

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 28, 2002, 11:27:19 ; Search time 21.78 seconds

(without alignments)  
405.329 Million cell updates/sec

Title: US-09-853-526-70

Perfect score: 228  
Sequence: 1 MKYLLPSVVLGTAPYVLA.....NYLDAIVDTVVEGKDDGG 228

Scoring table: OLIGO  
Gapop 60.0 , Gapext 60.0

Searched: 105224 seqs, 38719550 residues

Word size : 0

Total number of hits satisfying chosen parameters: 105224

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database : SwissProt\_40.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	228	100.0	353	1 PICE_HUMAN	Q9nqg2 homo sapien
2	32	14.0	354	1 PICE_MOUSE	Q9d1e8 mus musculu
3	9	3.9	237	1 PISC_HELPY	Q92jnh mus musculu
4	9	3.9	240	1 PISC_HELPY	Q25903 helicobacte
5	8	3.5	396	1 YD18_YEAST	Q12188 helicobacte
6	8	3.5	611	1 HSCA_BUCAT	P57660 buchneia ap
7	7	3.1	57	1 YDJO_ECOLI	P58038 escherichia
8	7	3.1	194	1 PURI_LACCA	P35853 lactobacill
9	7	3.1	208	1 HBGF_CERAE	Q09118 cercopithec
10	7	3.1	208	1 HBGF_HUMAN	Q09175 homo sapien
11	7	3.1	208	1 HBGF_HUMAN	Q01580 sus scrofa
12	7	3.1	250	1 PLSC_BORBU	Q06175 rattus norv
13	7	3.1	250	1 PLSC_BORBU	Q05188 borrelia bu
14	7	3.1	286	1 PSD8_CAEEL	Q23449 caenorhabdi
15	7	3.1	286	1 PSD8_CAEEL	P32263 saccharomyc
16	7	3.1	308	1 PROC_YEAST	Q42670 coccos nucif
17	7	3.1	445	1 XYLA_BACSU	P13860 phenocochae
18	7	3.1	516	1 GUX1_PHACU	P10435 nocardia la
19	7	3.1	520	1 CMCH_NOCUA	O51080 nocardia la
20	7	3.1	532	1 GSI_NEUCR	P36678 neurospora
21	7	3.1	677	1 Y593_TREPA	O81602 treponema p
22	7	3.1	860	1 CH12_COCIM	P54197 coccioidoide
23	7	3.1	904	1 Y002_CAEEL	Q09228 caenorhabdi
24	7	3.1	904	1 Y002_CAEEL	Q09228 caenorhabdi
25	6	2.6	25	1 PCWL_HUMAN	P12923 homo sapien
26	6	2.6	25	1 PCWL_PACGO	P82423 pachycondyl
27	6	2.6	76	1 PCWL_PACGO	P82424 pachycondyl
28	6	2.6	83	1 BB11_SCHCO	P78742 schizophy11
29	6	2.6	87	1 RL23_HALGA	Q06842 halobacteri
30	6	2.6	88	1 YLMC_BPPH1	P10435 bacteriophi
31	6	2.6	88	1 YLMC_BPPH1	P10435 bacteriophi
32	6	2.6	99	1 YVOL_RHISN	P55597 rhizobium s
33	6	2.6	101	1 THST_THERS	O57832 methanococc

## ALIGNMENTS

RESULT	1	STANDARD:	PRT:	353 AA.
AC	09N02: Q9B0G4:			
AC	09N02: Q9B0G4:			
DT	01-MAR-2002 (Rel. 41, Created)			
DT	01-MAR-2002 (Rel. 41, Last sequence update)			
DT	01-MAR-2002 (Rel. 41, Last annotation update)			
DE	1-acyl-sn-glycerol-3-phosphate acyltransferase epsilon (EC 2.3.1.51)			
DE	(1-AGP acyltransferase 5) (1-AGPAT 5) (lysophosphatidic acid			
DE	acyltransferase-epsilon) (LPAAT-epsilon) (1-acylglycerol-3-phosphate			
DE	O-acyltransferase 5).			
GN	AGPAT5.			
GN	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Euteria; Primates; Catarrhini; Hominidae; Homo.			
OX	NCBI_Taxid-9606;			
RP	SEQUENCE FROM N.A.			
RA	Leung D.W.;			
RT	"Cloning and expression of lPAAT-epsilon."			
RT	Submitted (MAY-2001) to the EMBL/Genbank/DBJ databases.			
RP	SEQUENCE FROM N.A.			
RA	Cohen D., Chumakov I., Blumenfeld M., Bougueleret L.;			
RA	Patent number WO9932644, 01-JUL-1999.			
RP	SEQUENCE FROM N.A.			
RA	MEDLINE-21154917; PubMed-11230166;			
RA	Wiemann S., Weil B., Wellenreuther R., Gassenhuber J., Glasel S.,			
RA	Ansorge W., Boecker M., Bloeker H., Bauersachs S., Blum H.,			
RA	Lauber J., Duesterhoeft A., Beyer A., Koehler K., Strack N.,			
RA	Mewes H.-W., Oltmannseder B., Obermaier B., Tampe J., Heubner D.,			
RA	Wandut R., Korn B., Klein M., Poustka A.;			
RT	"Towards a catalog of human genes and proteins: sequencing and			
RT	analysis of 500 novel complete protein coding human cDNAs."			
RT	Genome Res. 11:422-435(2001).			
RP	SEQUENCE FROM N.A.			
RA	Tissue-Placenta:			
RA	Isogai T., Oca T., Hayashi K., Sugiyama T., Otsuki T., Suzuki Y.,			
RA	Nishikawa T., Nagai K., Sugano S., Shitatori A., Sudo H.,			
RA	Yagatsuma M., Hosoliri T., Kaku Y., Kodaira H., Kondo H., Sugawara M.,			
RA	Takahashi S., Chiba Y., Ishida S., Murakami K., Ono Y., Takiguchi S.,			
RA	Watanabe S., Kimura K., Murakami K., Ishii S., Kawai Y., Saito K.,			
RA	Yamanoto J., Wakamatsu A., Nakamura Y., Nishihara K., Masuno Y.,			
RA	Nimomiya K., Iwyanagi T.;			
RT	"NDO human cDNA sequencing project."			
RT	Submitted (FEB-2000) to the EMBL/Genbank/DBJ databases.			
RP	FUNCTION: CONVERTS LYSOPHOSPHATIDIC ACID (LPA) INTO PHOSPHATIDIC			
RP	ACID BY INCORPORATING ACYL MOIETY AT THE 2 POSITION (BY			
RP	STIMULATORY).			
RP	CAATALYTIC ACTIVITY: Acyl-CoA + 1-acyl-sn-glycerol 3-phosphate =			
RP	CoA + 1,2-diacyl-sn-glycerol 3-phosphate.			
RP	PATHWAY: SECOND STEP IN DE NOVO PHOSPHOLIPID BIOSYNTHESIS.			
RP	-1- SUBCELLULAR LOCATION: Integral membrane protein (Potential).			

CC -1- SIMILARITY: BELONGS TO THE 1-ACYL-SN-GLYCEROL-3-PHOSPHATE  
CC ACYLTRANSFERASE FAMILY.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation  
CC at the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
CC -----  
CC EMBL: AF375789; AAK54809.1; ALT\_INT.  
CC EMBL: AL136587; CAB66522.1; ALT\_INT.  
CC EMBL: AK002072; BAA32069.1; -  
CC InterPro: IPR002123; Acyltransferase.  
CC Pfam: PF01553; Acyltransferase: 1.  
CC Phospholipid biosynthesis; Transferase; Acyltransferase;  
CC Transmembrane.  
CC TRANSMEM 7 29 POTENTIAL.  
CC TRANSMEM 44 56 POTENTIAL.  
CC TRANSMEM 334 351 POTENTIAL.  
CC CONFLICT 145 145 L -> V (IN REF. 2).  
CC SQUENCE 353 AA; 40813 MW; A05B1FA246CE1B64 CRC64;  
  
Query Match 100.0%; Score 228; DB 1; Length 353;  
Best Local Similarity 100.0%; Pred. No. 1.4e-227;  
Matches 228; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
OY 1 MRYLLPSVLLGTAPYVYLANGVRLSAPLPARFYQALDDRLCYQSNVLEFFENYTG 60  
DB 1 MRYLLPSVLLGTAPYVYLANGVRLSAPLPARFYQALDDRLCYQSNVLEFFENYTG 60  
OY 61 VOILLYDDLPKKNKNIYILANHOSTVDMYADIAIRQNALGHRYVYLGKGLKMLFYYGC 120  
DB 61 VOILLYDDLPKKNKNIYILANHOSTVDMYADIAIRQNALGHRYVYLGKGLKMLFYYGC 120  
OY 121 YFAOHGCIYVRSKAFENEKEMRNKLSQSYVDAGTPMYLVIEPEGRYVPEQTKVLSAQAF 180  
DB 121 YFAOHGCIYVRSKAFENEKEMRNKLSQSYVDAGTPMYLVIEPEGRYVPEQTKVLSAQAF 180  
OY 181 AAOGLAVLKHVLPRIKATHVAFDCKMKNYLDATYDTVYVEGKDDGG 228  
DB 181 AAOGLAVLKHVLPRIKATHVAFDCKMKNYLDATYDTVYVEGKDDGG 228  
  
RESULT 2  
PLSC\_MOUSE STANDARD; PRT; 354 AA.  
AC Q9D1E8;  
DT 01-MAR-2002 (Rel. 41, Created)  
DT 01-MAR-2002 (Rel. 41, Last sequence update)  
DT 01-MAR-2002 (Rel. 41, Last annotation update)  
DE 1-acyl-sn-glycerol-3-phosphate acyltransferase epsilon (EC 2.3.1.51)  
DE (1-AGP acyltransferase 5) (1-AGPAT 5) (lysophosphatidic acid  
DE acyltransferase-epsilon) (lPAAT-epsilon) (1-acylglycerol-3-phosphate  
DE O-acyltransferase 5).  
GN AGPAT5.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCBI\_TaxID=10090;  
RN [1]  
RP SEQUENCE FROM N. A.  
RC STRAIN=C57BL/6J; TISSUE=Embryo;  
RX MEDLINE=21085660; PubMed=11217851;  
RA Araiwa T., Hara A., Fukunishi Y., Kono H., Adachi J., Fukuda S.,  
RA Araiwa K., Iwawa N., Nishi K., Kiyosawa H., Kondo S., Yamana K. I.,  
RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,  
RA Kadota K., Matsuda H. A., Ashburner M., Batilov S., Casavant T.,  
RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,  
RA Knehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,

RA Schirml L. M., Staabli F., Suzuki R., Tomita M., Wagner L., Washio T.,  
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,  
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M. F.,  
RA Brownstein M. J., Bult C., Fletcher C., Fujita M., Gariboldi M.,  
RA Cusumini S., Hill D., Hofmann M., Hume D. A., Kamiya M., Lee N. H.,  
RA Lyons F., Marchionni L., Mashima J., Mazzarelli J., Nombarts P.,  
RA Norioka P., Ring B., Ringwald M., Rodriguez J., Sakamoto N.,  
RA Sasaki H., Sato K., Schoenbach C., Seva T., Shibata Y., Storch K. F.,  
RA Suzuki H., Toyooka K., Wang K. H., Weltz C., Whitaker C., Wilming L.,  
RA Wyszewski-Boris A., Yoshida K., Hasegawa Y., Kawai H., Kohsaki S.,  
RA Hayashizaki Y.;  
RA "Functional annotation of a full-length mouse cDNA collection";  
RL Nature 403:680-690(2001).  
CC -1- FUNCTION: CONVERTS LYOPHOSPHATIDIC ACID (LPA) INTO PHOSPHATIDIC  
CC ACID BY INCORPORATING ACYL MOIETY AT THE 2 POSITION (BY  
CC SIMILARITY).  
CC -1- CATALYTIC ACTIVITY: Acyl-CoA + 1-acyl-sn-glycerol 3-phosphate -  
CC CoA + 1,2-diacyl-sn-glycerol 3-phosphate.  
CC -1- PATHWAY: SECOND STEP IN DE NOVO PHOSPHOLIPID BIOSYNTHESIS.  
CC -1- SUBCELLULAR LOCATION: Integral membrane protein (Potential).  
CC -1- SIMILARITY: BELONGS TO THE 1-ACYL-SN-GLYCEROL-3-PHOSPHATE  
CC ACYLTRANSFERASE FAMILY.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation.  
CC at the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
CC -----  
CC EMBL: AK006449; BAB22915.1; -  
CC MGD; MGI:1915880; Aspats.  
CC DR InterPro: IPR002123; Acyltransferase.  
CC DR Pfam: PF01553; Acyltransferase: 1.  
CC KW Phospholipid biosynthesis; Transferase; Acyltransferase;  
CC Transmembrane.  
CC TRANSMEM 7 29 POTENTIAL.  
CC TRANSMEM 44 66 POTENTIAL.  
CC TRANSMEM 335 352 POTENTIAL.  
CC SQUENCE 354 AA; 40943 MW; 447EEF924B91E800 CRC64;  
  
Query Match 14.0%; Score 32; DB 1; Length 354;  
Best Local Similarity 100.0%; Pred. No. 3e-25;  
Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
OY 174 LSASQAFPAORGVLAVLKHVLPRIKATHVAFD 205  
DB 174 LSASQAFPAORGVLAVLKHVLPRIKATHVAFD 205  
  
RESULT 3  
PLSC\_HELPJ STANDARD; PRT; 237 AA.  
AC Q9ZJN8;  
DT 16-OCT-2001 (Rel. 40, Created)  
DT 16-OCT-2001 (Rel. 40, Last sequence update)  
DT 16-OCT-2001 (Rel. 40, Last annotation update)  
DE 1-acyl-sn-glycerol-3-phosphate acyltransferase (EC 2.3.1.51) (1-AGP  
DE acyltransferase) (1-AGPAT) (lysophosphatidic acid acyltransferase)  
DE (LPAAT).  
GN PLSC OR JHP1267.  
OS Helicobacter pylori 399 (Campylobacter pylori 399).  
OC Bacteria; Proteobacteria; epsilon subdivision; Helicobacter group;  
OC Helicobacter.  
OX NCBI\_TaxID=85963;  
RN [1]  
RP SEQUENCE FROM N. A.  
RX MEDLINE=99120557; PubMed=9923682;  
RA Alm R. A., Ling L. S. L., Moid D. T., King B. L., Brown E. D., Doig P. C.,  
RA Smith D. R., Noonan B., Guild B. C., deJonge B. L., Carmel G.,  
RA Tummino P. J., Caruso A., Uria-Nickelsen M., Mills D. M., Ives C.,



RA Gibson R., Merberg D., Mills S.D., Jiang Q., Taylor D.E., Voyis G.F.,  
 RA Trust T.J.;  
 RT "Genomic sequence comparison of two unrelated isolates of the human  
 RT gastric pathogen *Helicobacter pylori*.";  
 RL Nature 397:176-180(1999).  
 CC -1- FUNCTION: CONVERTS LYOPHOSPHATIDIC ACID (LPA) INTO PHOSPHATIDIC  
 CC ACID BY INCORPORATING ACYL MOIETY AT THE 2 POSITION.  
 CC -1- CATALYTIC ACTIVITY: Acyl-CoA + 1-acyl-sn-glycerol 3-phosphate =  
 CC CoA + 1,2-diacyl-sn-glycerol 3-phosphate.  
 CC -1- PATHWAY: SECOND STEP IN DE NOVO PHOSPHOLIPID BIOSYNTHESIS.  
 CC -1- SUBCELLULAR LOCATION: Inner membrane-associated (Potential).  
 CC -1- SIMILARITY: BELONGS TO THE 1-ACYL-SN-GLYCEROL-3-PHOSPHATE  
 CC ACYLTRANSFERASE FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL: AE001550; AAD06852.1; -  
 CC InterPro: IPR002123; Acyltransferase.  
 CC Pfam: PF01553; Acyltransferase; 1.  
 CC KMW Phospholipid biosynthesis; Transferase; Acyltransferase;  
 CC KMW Inner membrane; Complete proteome  
 CC KMW SEQUENCE 237 AA; 27188 MW; E10F517D42A1731F CRC64;

Query Match 3.9%; Score 9; DB 1; Length 237;  
 Best Local Similarity 100.0%; Pred. No. 0.12;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 157 LVIFPEGTR 165  
 |||||||||  
 DB 142 LVIFPEGTR 150

RESULT 4  
 PLSC\_HELPY STANDARD: PRT; 240 AA.  
 AC 025903;  
 DT 15-DEC-1998 (Rel. 37, Last sequence update)  
 DT 15-DEC-1998 (Rel. 37, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE 1-acyl-sn-glycerol-3-phosphate acyltransferase (EC 2.3.1.51) (1-AGP  
 DE acyltransferase) (1-AGPAT) (Lyso-phosphatidic acid acyltransferase)  
 DE (LPAAT).  
 GN PLSC OR HP1348.  
 OS *Helicobacter pylori* (Campylobacter *pylori*).  
 OC Bacteria; Proteobacteria; epsilon subdivision; *Helicobacter* group;  
 OC *Helicobacter*.  
 OX NCBI\_TaxID=210;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=26695 / ATCC 700392;  
 RX MEDLINE=97394467; PubMed=9252185;  
 RA Toml J.-F., White O., Kerlavage A.R., Clayton R.A., Sutton G.G.,  
 RA Fleischmann R.D., Ketchum K.A., Klenk H.-P., Gill S., Dougherty B.A.,  
 RA Nelson K., Quackenbush J., Zhou L., Kirkness E.F., Peterson S.,  
 RA Loftus B., Fitzgerald D., Dodson R., Khaliq H.G., Glodek A.,  
 RA McKenney K., Fitzgerald L.M., Lee N., Adams M.D., Hickey E.K.,  
 RA Berg D.E., Gocayne J.D., Utterback T.R., Peterson J.D., Kelley J.M.,  
 RA Cotton M.D., Weidman J.M., Fujii C., Bowman C., Watthey L., Wallin E.,  
 RA Hayes W.S., Borodovsky M., Karp P.D., Smith H.O., Fraser C.M.,  
 RA Venter J.C.;  
 RT "The complete genome sequence of the gastric pathogen *Helicobacter*  
 RT *pylori*.";  
 RL Nature 388:539-547(1997).  
 CC -1- FUNCTION: CONVERTS LYOPHOSPHATIDIC ACID (LPA) INTO PHOSPHATIDIC  
 CC ACID BY INCORPORATING ACYL MOIETY AT THE 2 POSITION.  
 CC -1- CATALYTIC ACTIVITY: Acyl-CoA + 1-acyl-sn-glycerol 3-phosphate =

CC CoA + 1,2-diacyl-sn-glycerol 3-phosphate.  
 CC -1- PATHWAY: SECOND STEP IN DE NOVO PHOSPHOLIPID BIOSYNTHESIS.  
 CC -1- SUBCELLULAR LOCATION: Inner membrane-associated (Potential).  
 CC -1- SIMILARITY: BELONGS TO THE 1-ACYL-SN-GLYCEROL-3-PHOSPHATE  
 CC ACYLTRANSFERASE FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL: AE000636; AAD08393.1; -  
 CC TIGR: HP1348; -  
 CC InterPro: IPR002123; Acyltransferase.  
 CC Pfam: PF01553; Acyltransferase; 1.  
 CC KMW Phospholipid biosynthesis; Transferase; Acyltransferase;  
 CC KMW Inner membrane; Complete proteome  
 CC KMW SEQUENCE 240 AA; 27745 MW; 22BD50DEB190BBD CRC64;

Query Match 3.9%; Score 9; DB 1; Length 240;  
 Best Local Similarity 100.0%; Pred. No. 0.12;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 157 LVIFPEGTR 165  
 |||||||||  
 DB 142 LVIFPEGTR 150

RESULT 5  
 YD18\_YEAST STANDARD: PRT; 396 AA.  
 AC 012185;  
 DT 15-JUL-1999 (Rel. 38, Created)  
 DT 15-JUL-1999 (Rel. 38, Last sequence update)  
 DT 15-JUL-1999 (Rel. 38, Last annotation update)  
 DE Hypothetical 45.9 kDa protein in KCS1-GCV1 intergenic region.  
 GN YDR018C OR YD9335.04C OR P2F396.  
 OS *Saccharomyces cerevisiae* (Baker's yeast).  
 CC Eukaryota; Fungi; Ascomycota; *Saccharomycotina*; *Saccharomycetes*;  
 CC *Saccharomycetales*; *Saccharomycetaceae*; *Saccharomyces*.  
 OX NCBI\_TaxID=4932;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=S288C / AB972;  
 RA Dedman K., Brown D., Hamlyn N., Bowman S., Barrell B.G.,  
 RA Rajandream M.A.;  
 RL Submitted (MAY-1995) to the EMBL/Genbank/DBJ databases.  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=97051598; PubMed=8896275;  
 RA Elde L.G., Sander C., Prydz H.;  
 RT "Sequencing and analysis of a 35.4 kb region on the right arm of  
 RT chromosome IV from *Saccharomyces cerevisiae* reveal 23 open reading  
 RT frames.";  
 RL Yeast 12:1085-1090(1996).  
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein (Potential).  
 CC -1- SIMILARITY: BELONGS TO THE 1-ACYL-SN-GLYCEROL-3-PHOSPHATE  
 CC ACYLTRANSFERASE FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL: 274314; CAA98838.1; -  
 CC EMBL: X95966; CAA65210.1; -

DR EMBL: 249770; CAAB9843.1; -.  
 DR SGD: S0002425; YDR018C.  
 DR InterPro: IPR002123; Acyltransferase.  
 DR Pfam: PF01553; Acyltransferase; 1.  
 KM Hypothetical protein: Phospholipid biosynthesis; Transferase;  
 KM Acyltransferase; Transmembrane.  
 FT TRANSMEM 27 4/7 POTENTIAL.  
 FT TRANSMEM 69 6/9 POTENTIAL.  
 FT TRANSMEM 123 143 POTENTIAL.  
 FT TRANSMEM 372 392 POTENTIAL.  
 SO SEQUENCE 396 AA; 45938 MW; 9F55AB56C8FD44DD CRC64;

Query Match 3.5%; Score 8; DB 1; Length 396;  
 Best Local Similarity 100.0%; Pred. No. 1.9;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 211 LDAIYDVT 218  
 DB 265 LDAIYDVT 272

RESULT 6  
 HSCA\_BUCAI STANDARD; PRT; 611 AA.  
 ID HSCA\_BUCAI STANDARD; PRT; 611 AA.  
 AC P57660;  
 DT 16-OCT-2001 (Rel. 40, Created)  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Chaperone protein hscA homolog.  
 GN HSCA OR BU605.  
 OS Buchnera aphidicola (subsp. Acyrthosiphon pisum) (Acyrthosiphon pisum  
 symbiotic bacterium).  
 OC Bacteria; Proteobacteria; gamma subdivision; Buchnera.  
 OX NCBI\_TaxID=118099;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=Tokyo 1998;  
 RX MEDLINE=20445173; PubMed=10993077;  
 RA Shigenobu S., Watanabe H., Hattori M., Sakaki Y., Ishikawa H.;  
 RT "Genome sequence of the endocellular bacterial symbiont of aphids  
 Buchnera sp. AS.";  
 RL Nature 407:81-86(2000).  
 CC -1- FUNCTION: PROBABLE CHAPERONE. HAS A LOW INTRINSIC ATPASE ACTIVITY  
 WHICH IS MARKEDLY STIMULATED BY HSCB (BY SIMILARITY).  
 CC -1- SIMILARITY: BELONGS TO THE HEAT SHOCK PROTEIN 70 FAMILY.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 CC EMBL: AP001119; BAB13289.1; -.  
 DR InterPro: IPR01023; HSP70.  
 DR Pfam: PF00012; HSP70; 1.  
 DR PRINTS: PR00301; HEATSHOCK70.  
 DR PROSITE: PS00287; HSP70\_1; 1.  
 DR PROSITE: PS00329; HSP70\_2; 1.  
 DR PROSITE: PS01036; HSP70\_3; 1.  
 KM Chaperone; ATP-binding; Complete proteome.  
 SO SEQUENCE 611 AA; 69837 MW; 32B567C53073082A CRC64;

Query Match 3.5%; Score 8; DB 1; Length 611;  
 Best Local Similarity 100.0%; Pred. No. 2.8;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 RYLLPSV 9  
 DB 42 RYLLPSV 49

RESULT 7  
 YJDO\_ECOLI STANDARD; PRT; 57 AA.  
 ID YJDO\_ECOLI STANDARD; PRT; 57 AA.  
 AC P58038;  
 DT 16-OCT-2001 (Rel. 40, Created)  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DT 01-MAR-2002 (Rel. 41, Last annotation update)  
 DE Hypothetical protein YJDO.  
 GN YJDO OR B4128.1 OR 25731 OR EC55110.1.  
 OS Escherichia coli, and  
 OS Escherichia coli O157:H7.  
 OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;  
 OC Escherichia.  
 OX NCBI\_TaxID=562, 83334;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=K12 / MG1655;  
 RX MEDLINE=97426617; PubMed=9278503;  
 RA Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,  
 RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,  
 RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,  
 Mau B., Shao Y.;  
 RT "The complete genome sequence of Escherichia coli K-12.";  
 RL Science 277:1453-1474(1997).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=O157:H7 / EDL933 / ATCC 700927;  
 RX MEDLINE=21074935; PubMed=11206551;  
 RA Perna N.T., Plunkett G. III, Burland V., Mau B., Glasner J.D.,  
 RA Rose D.J., Mayhew G.F., Evans P.S., Gregor J., Kirkpatrick H.A.,  
 RA Postel E.J., Hackett J., Klink S., Boutin A., Shao Y., Miller L.,  
 RA Grobeck E.J., Davis N.W., Lim A., Diallanla E.T., Potamousis K.,  
 RA Apodaca J., Anantharaman T.S., Lin J., Yen G., Schwartz D.C.,  
 RA Welch R.A., Blattner F.R.;  
 RT "Genome sequence of enterohaemorrhagic Escherichia coli O157:H7.";  
 RL Nature 409:529-533(2001).  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=O157:H7 / RIMD 0509952;  
 RX MEDLINE=21156231; PubMed=11258796;  
 RA Hayashi T., Makino K., Onishi M., Kurokawa K., Ishii K., Yokoyama K.,  
 RA Han C.-G., Ohtsubo E., Nakayama K., Murata T., Tanaka M., Tobe T.,  
 RA Iida T., Takami H., Honda T., Sasakawa C., Ogasawara N., Yasunaga T.,  
 RA Kihara S., Shiba T., Hattori M., Shingawa H.;  
 RT "Complete genome sequence of enterohaemorrhagic Escherichia coli  
 O157:H7 and genomic comparison with a laboratory strain K-12.";  
 RL DNA Res. 8:11-22(2001).  
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein (Potential).  
 CC -1- SIMILARITY: STRONG, TO E.COLI YDCX.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 CC EMBL: AE000485; -; NOT\_ANNOTATED\_CDS.  
 DR EMBL: AE005646; AAG59328.1; -.  
 DR EMBL: AP002568; -; NOT\_ANNOTATED\_CDS.  
 DR Ecogene: BG14342; YJDO.  
 KM Hypothetical protein: Transmembrane; Complete proteome.  
 SO SEQUENCE 57 AA; 6555 MW; A3670A19500F75D6 CRC64;

Query Match 3.1%; Score 7; DB 1; Length 57;  
 Best Local Similarity 100.0%; Pred. No. 4;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 25 RLISAF 31  
 DB 33 RLISAF 39

RESULT 8  
 ID PURI\_LACCA STANDARD; PRT: 194 AA.  
 AC P35853;  
 DT 01-FEB-1994 (Rel. 29, Created)  
 DT 01-FEB-1996 (Rel. 33, Last sequence update)  
 DT 30-MAY-2000 (Rel. 39, Last annotation update)  
 DE Amidophosphoribosyltransferase precursor (EC 2.4.2.14) (Glutamine  
 phosphoribosylpyrophosphate amidotransferase) (ATPase) (GATase)  
 DE (Fragment).  
 GN PURP.  
 OS Lactobacillus casei.  
 OC Bacteria; Firmicutes; Bacillus/Clostridium group; Lactobacillaceae;  
 OC Lactobacillus.  
 OX NCBI\_TaxID=1582;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=93012962; PubMed=1398079;  
 RA Gu Z.-M., Martindale D.W., Lee B.H.;  
 RT "Isolation and complete sequence of the purL gene encoding PCAM  
 synthase II in Lactobacillus casei.";  
 RL Gene 119:123-126(1992).  
 RN [2]  
 RP ERRATUM.  
 RX MEDLINE=94040790; PubMed=8224889;  
 RA Gu Z.-M., Martindale D.W., Lee B.H.;  
 RL Gene 133:147-147(1993).  
 CC -1- CATALYTIC ACTIVITY: 5-phospho-beta-D-ribosylamine + diphosphate +  
 L-glutamate -> L-glutamine + 5-phospho-alpha-D-ribose 1-diphosphate  
 + H(2O).  
 CC -1- PATHWAY: FIRST STEP IN DE NOVO PURINE BIOSYNTHESIS.  
 CC -1- SIMILARITY: THE GATASE DOMAIN BELONGS TO TYPE-2 GLUTAMINE  
 AMIDOTRANSFERASES.  
 CC -1- SIMILARITY: IN THE C-TERMINAL SECTION, BELONGS TO THE  
 PURINE/PRIMIDINE PHOSPHORIBOSYLTRANSFERASE FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation  
 at the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL: M85265; AAC36948.1; .  
 DR PIR: PC1136; PC1136.  
 DR HSSP: P00497; 1A00.  
 DR MEROPS: C44.001; .  
 DR InterPro: IPR000583; GATase\_2.  
 DR InterPro: IPR002375; Pur\_pyr\_pr\_transf.  
 DR Pfam: PF00310; GATase\_2; 1.  
 DR PROSITE: PS00103; PUR\_PYR\_PR\_TRANSFER; PARTIAL.  
 DR PROSITE: PS00443; GATASE\_TYPE\_II; 1.  
 KW Purine biosynthesis; Transferase; Glycosyltransferase.  
 FT PROPEP 1 11 BY SIMILARITY.  
 FT CHAIN 12 >194 AMIDOPHOSPHORIBOSYLTRANSFERASE.  
 FT ACT\_SITE 12 12 GATASE (INVOLVED IN SUBSTRATE BINDING AND  
 CATALYSIS) (BY SIMILARITY).  
 FT NON\_TER 194 194  
 SO SEQUENCE 194 AA; 21144 MW; 4A78BCB5365D5EC CRC64;

Query Match 3.1%; Score 7; DB 1; Length 194;  
 Best Local Similarity 100.0%; Pred. No. 11;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 100 ALGHVRY 106

DB 79 ALGHVRY 85

RESULT 9  
 ID HBGF\_CERAE STANDARD; PRT: 208 AA.  
 AC 009118;  
 DT 01-FEB-1996 (Rel. 33, Created)  
 DT 01-FEB-1996 (Rel. 33, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Heparin-binding EGF-like growth factor precursor (HB-EGF) (HBEGF)  
 DE (Diphtheria toxin receptor) (DT-R).  
 GN DTR OR HBGF.  
 OS Cercopithecus aethiops (Green monkey) (Givet).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecidae;  
 OC Cercopithecinae; Cercopithecus.  
 OX NCBI\_TaxID=9534;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=92298386; PubMed=1606612;  
 RA Naglich J.G., Metherall J.E., Russel D.W., Bidelis L.;  
 RT "Expression cloning of a diphtheria toxin receptor: Identity with a  
 heparin-binding EGF-like growth factor precursor.";  
 RL Cell 69:1051-1061(1992).  
 RN [2]  
 RP TOXIN-BINDING DOMAIN.  
 RX MEDLINE=95126975; PubMed=7826391;  
 RA Hooper K.P., Bidelis L.;  
 RT "Localization of a critical diphtheria toxin-binding domain to the C-  
 terminus of the mature heparin-binding EGF-like growth factor region  
 of the diphtheria toxin receptor.";  
 RL Biochem. Biophys. Res. Commun. 206:710-717(1995).  
 CC -1- FUNCTION: MAY BE INVOLVED IN MACROPHAGE-MEDIATED CELLULAR  
 PROLIFERATION. IT IS MITOGENIC FOR FIBROBLASTS AND SMOOTH MUSCLE  
 BUT NOT ENDOTHELIAL CELLS. IT IS ABLE TO BIND EGF RECEPTORS WITH  
 HIGHER AFFINITY THAN EGF ITSELF AND IS A FAR MORE POTENT MITOGEN  
 FOR SMOOTH MUSCLE CELLS THAN EGF (BY SIMILARITY).  
 CC -1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN. MATURE HB-EGF IS  
 RELEASED INTO THE EXTRACELLULAR SPACE AND PROBABLY BINDS TO A  
 RECEPTOR (BY SIMILARITY).  
 CC -1- PTM: O-GLYCOSYLATED (BY SIMILARITY).  
 CC -1- SIMILARITY: CONTAINS 1 EGF-LIKE DOMAIN.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation  
 at the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL: M93012; .; NOT\_ANNOTATED\_CD5.  
 DR HSSP: Q99075; 1XDT.  
 DR InterPro: IPR000561; EGF-like.  
 DR Pfam: PF00008; EGF; 1.  
 DR SMART: SM00181; EGF; 1.  
 DR PROSITE: PS00022; EGF\_1; 1.  
 DR PROSITE: PS01186; EGF\_2; 1.  
 KW Growth factor; Heparin-binding; Signal; EGF-like domain;  
 KW Transmembrane; Glycoprotein; Receptor.  
 FT SIGNAL 1 19  
 FT PROPEP 20 62  
 FT CHAIN 63 148  
 FT PROPEP 149 208  
 FT DOMAIN 20 160  
 FT DOMAIN 161 184  
 FT TRANSMEM 185 208  
 FT DOMAIN 104 144  
 FT DOMAIN 104 144  
 FT CARBOHYD 75 75  
 FT CARBOHYD 85 85  
 FT DISULFID 108 121  
 BY SIMILARITY.  
 HEPARIN-BINDING EGF-LIKE GROWTH FACTOR.  
 C-TERMINAL (POTENTIAL).  
 EXTRACELLULAR (POTENTIAL).  
 POTENTIAL.  
 CYTOPLASMIC (POTENTIAL).  
 EGF-LIKE.  
 O-LINKED (GALNAC. . .) (BY SIMILARITY).  
 O-LINKED (GALNAC. . .) (BY SIMILARITY).  
 BY SIMILARITY.

```

FT DISULFID 116 132 BY SIMILARITY.
FT DISULFID 134 143 BY SIMILARITY.
SQ SEQUENCE 208 AA: 22985 MW: 8D108289AD0485AE9 CRC64:

Query Match 3.1%; Score 7; DB 1; Length 208;
Best Local Similarity 100.0%; Pred. No. 12;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 4 LPSVVL 10
Db 3 LPSVVL 9

RESULT 10
HBGF_HUMAN STANDARD: PRT: 208 AA.
AC 099075;
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Heparin-binding EGF-like growth factor precursor (HB-EGF) (HBEGF)
DE (Diphtheria toxin receptor) (DTR-R).
GN DTR OR HBGF.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A., AND SEQUENCE OF 73-93.
RC TISSUE=Macrophage;
RX MEDLINE=91157008; PubMed=1840698;
RA Higashiyama S., Abraham J.A., Miller J., Fiddes J.C., Klagsbrun M.;
RT "A heparin-binding growth factor secreted by macrophage-like cells
RT that is related to EGF.";
RL Science 251:936-939(1991).
RN [2]
RP SEQUENCE FROM N.A.
RA Kimmery W., Bondoc M., Cheng J., Connolly K.S., Gunning K.M.,
RA Kader K., Miguel T., Miller C., Pittluck S., Pollard M., Rojeski H.,
RA Subramanian S., Martin C.H.;
RL Submitted (APR-1998) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE OF 63-141 AND 143-148, AND CARBOHYDRATE-LINKAGE SITES.
RC TISSUE=Histocytic Lymphoma;
RX MEDLINE=92210596; PubMed=1556128;
RA Higashiyama S., Lau K., Besner G.E., Abraham J.A., Klagsbrun M.;
RT "Structure of heparin-binding EGF-like growth factor. Multiple forms,
RT primary structure, and glycosylation of the mature protein.";
RL J. Biol. Chem. 267:6205-6212(1992).
RN [4]
RP TOXIN-BINDING DOMAIN.
RX MEDLINE=95138082; PubMed=7836353;
RA Mitamura T., Higashiyama S., Taniguchi N., Klagsbrun M., Mekada E.;
RT "Diphtheria toxin binds to the epidermal growth factor (EGF)-like
RT domain of human heparin-binding EGF-like growth factor/diphtheria
RT toxin receptor and inhibits specifically its mitogenic activity.";
RL J. Biol. Chem. 270:1015-1019(1995).
RN [5]
RP X-RAY CRYSTALLOGRAPHY (2.65 ANGSTROMS) OF 73-147 IN COMPLEX WITH TOX.
RX MEDLINE=98324089; PubMed=9559904;
RA Louie G.V., Yang W., Bowman M.E., Choe S.;
RT "Crystal structure of the complex of diphtheria toxin with an
RT extracellular fragment of its receptor.";
RL Mol. Cell 1:67-78(1997).
CC -1- FUNCTION: MAY BE INVOLVED IN MACROPHAGE-MEDIATED CELLULAR
CC PROLIFERATION. IT IS MITOGENIC FOR FIBROBLASTS AND SMOOTH MUSCLE
CC BUT NOT ENDOTHELIAL CELLS. IT IS ABLE TO BIND EGF RECEPTORS WITH
CC HIGHER AFFINITY THAN EGF ITSELF AND IS A FAR MORE POTENT MITOGEN
CC FOR SMOOTH MUSCLE CELLS THAN EGF.
CC -1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN. MAJURE HB-EGF IS
CC RELEASED INTO THE EXTRACELLULAR SPACE AND PROBABLY BINDS TO A
CC RECEPTOR.

```

```

CC -1- PTM: SEVERAL N-TERMINI HAVE BEEN IDENTIFIED BY DIRECT SEQUENCING.
CC THE FORMS WITH N-TERMINI 63, 73 AND 74 HAVE BEEN TESTED AND FOUND
CC TO BE BIOLOGICALLY ACTIVE.
CC -1- PTM: O-GLYCAN ATTACHMENT SITES WERE DETERMINED BY EDMAN
CC DEGRADATION, O-GLYCANASE DIGESTS MUCIN-TYPE GLYCOSYLATION
CC (DONE IN HB-EGF FURIFIED FROM HISTOCYTIC LYMPHOMA CELL LINE
CC U-937).
CC -1- SIMILARITY: CONTAINS 1 EGF-LIKE DOMAIN.
CC -----
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
DR EMBL: M60278; AAA35956.1; -.
DR EMBL: AC004634; AAC15470.1; -.
DR PIR: A38432; A38432.
DR PDB: 1XDT; 25-FEB-98.
DR MIM: 126150; -.
DR InterPro: IPR000561; EGF-like.
DR Pfam: PF00008; EGF; 1.
DR SMART: SM00181; EGF; 1.
DR PROSITE: PS00022; EGF_1; 1.
DR PROSITE: PS01186; EGF_2; 1.
KW Growth factor; Heparin-binding; Signal; EGF-like domain;
KW Transmembrane; Glycoprotein; Receptor; 3D-structure.
FT SIGNAL 1 19
FT PROPEP 20 62
FT CHAIN 63 148
FT PROPEP 149 208
FT DOMAIN 20 160
FT TRANSMEM 161 184
FT DOMAIN 185 208
FT DOMAIN 104 144
FT DISULFID 108 121
FT DISULFID 116 132
FT DISULFID 134 143
FT CARBOHYD 75 75
FT CARBOHYD 85 85
SQ SEQUENCE 208 AA: 23067 MW: 2C43C9D1D8291B51 CRC64:

Query Match 3.1%; Score 7; DB 1; Length 208;
Best Local Similarity 100.0%; Pred. No. 12;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 4 LPSVVL 10
Db 3 LPSVVL 9

RESULT 11
HBGF_PIG STANDARD: PRT: 208 AA.
AC 001580;
DT 01-FEB-1996 (Rel. 33, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Heparin-binding EGF-like growth factor precursor (HB-EGF) (HBEGF).
GN DTR OR HBGF.
OS Sus scrofa (Pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX NCBI_TaxID=9823;
RN [1]
RP SEQUENCE FROM N.A.
RA Pascall J.C.;
RL Submitted (NOV-1997) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE OF 99-182 FROM N.A.

```

us-09-853-526-70.rsp

```

DE 16-OCT-2001 (Rel. 40, last annotation update)
DE Heparin-binding EGF-like growth factor precursor (HB-EGF) (HBEGF).
GN DTR OR HEGFL.
OS Rattus norvegicus (Rat).
CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
CC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=SPAGUE-DAWLEY; TISSUE=Macrophage;
RX MEDLINE=93135756; PubMed=7678488;
RA Abraham J.A., Damm D., Bajardi A., Miller J., Klagsbrun M.,
RA Ezekowitz R.A.B.;
RT "Heparin-binding EGF-like growth factor: characterization of rat and
RT mouse cDNA clones, protein domain conservation across species, and
RT transcript expression in tissues."
RL Biochem. Biophys. Res. Commun. 190:125-133(1993).
CC -1- FUNCTION: MAY BE INVOLVED IN MACROPHAGE-MEDIATED CELLULAR
CC PROLIFERATION. IT IS MITOGENIC FOR FIBROBLASTS AND SMOOTH MUSCLE
CC BUT NOT ENDOTHELIAL CELLS. IT IS ABLE TO BIND EGF RECEPTORS WITH
CC HIGHER AFFINITY THAN EGF ITSELF AND IS A FAR MORE POTENT MITOGEN
CC FOR SMOOTH MUSCLE CELLS THAN EGF.
CC -1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN. MATURE HB-EGF IS
CC RELEASED INTO THE EXTRACELLULAR SPACE AND PROBABLY BINDS TO A
CC RECEPTOR.
CC -1- TISSUE SPECIFICITY: MOST ABUNDANT IN SKELETAL MUSCLE, LUNG, SPLEEN
CC BRAIN AND HEART.
CC -1- PTM: O-GLYCOSYLATED (BY SIMILARITY).
CC -1- SIMILARITY: CONTAINS 1 EGF-LIKE DOMAIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; L05489; AAA81780.1; -.
DR PIR; JC1409; JC1409.
DR HSSP; 099075; 1XDT.
DR InterPro: IPR000561; EGF-like.
DR Pfam; PF00008; EGF; 1.
DR SMART; SM00181; EGF; 1.
DR PROSITE; PS00022; EGF_1; 1.
DR PROSITE; PS01186; EGF_2; 1.
KW Signal; Growth factor; Heparin-binding; EGF-like domain;
KW Transmembrane; Glycoprotein.
FT SIGNAL 1 23
FT PROPEP 24 62 POTENTIAL.
FT CHAIN 63 148 BY SIMILARITY.
FT PROPEP 149 208 HEPARIN-BINDING EGF-LIKE GROWTH FACTOR.
FT DOMAIN 24 160 C-TERMINAL (POTENTIAL).
FT TRANSLEM 161 184 EXTRACELLULAR (POTENTIAL).
FT DOMAIN 185 208 POTENTIAL.
FT DOMAIN 104 144 CYTOPLASMIC (POTENTIAL).
FT DISULFID 108 121 EGF-LIKE.
FT DISULFID 116 132 BY SIMILARITY.
FT DISULFID 134 143 BY SIMILARITY.
FT CARBOHYD 85 85 O-LINKED (GALNAC...) (BY SIMILARITY).
SQ SEQUENCE 208 AA; 22843 MW; DDBD045E116D064C CRC64;

Query Match 3.1%; Score 7; DB 1; Length 208;
Best Local Similarity 100.0%; Pred. No. 12;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 4 LLPSVVL 10
|111111|
Db 3 LLPSVVL 9

```

PLSC\_BORBU STANDARD; PRT: 250 AA.  
 ID PLSC\_BORBU  
 AC 059188;  
 DT 15-DEC-1998 (Rel. 37, Created)  
 DT 15-DEC-1998 (Rel. 37, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE 1-acyl-sn-glycerol-3-phosphate acyltransferase (EC 2.3.1.51) (1-AGP acyltransferase) (1-AGPAT) (lysophosphatidic acid acyltransferase) (LIPAT).  
 GN PLSC OR B80037.  
 OS Borrelia burgdorferi (Lyme disease spirochete).  
 OC Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.  
 OX NCBI\_TaxID=139;  
 RN (1)  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=ATCC 35210 / H31;  
 RX MEDLINE=98065943; PubMed=9403685;  
 RA Fraser C.M., Casjens S., Huang W.M., Sutton G.G., Clayton F.A., Lathigra R., White O., Ketchum K.A., Dodson R., Hickey E.K., Gwinn M., Dougherty B., Tomb J.-F., Fleischmann R.D., Richardson D., Peterson J., Kervatava A.R., Quackenbush J., Salzberg S., Hanson M., van Vugt R., Palmer N., Adams M.D., Gocayne J.D., Weidman C., Ulielback T., Matthey L., McDonald L., Artlisch P., Bowman C., Gierland S., Fujii C., Cotton M.D., Horst K., Roberts K., Hatch B., Smith H.O., Venter J.C.;  
 RA "Genomic sequence of a Lyme disease spirochete, Borrelia burgdorferi."  
 RT Nature 390:580-586(1997).  
 RL Nature 390:580-586(1997).  
 RN (2)  
 RP SEQUENCE OF 116-250 FROM N.A.  
 RC STRAIN=212;  
 RX MEDLINE=95111614; PubMed=7812434;  
 RA Ojalini C., Davidson B.E., Saint-Girons I., Old I.G.;  
 RA "Conservation of gene arrangement and an unusual organization of rRNA genes in the linear chromosomes of the Lyme disease spirochaetes Borrelia burgdorferi, B. garinii and B. afzelii."  
 RT Microbiology 140:2931-2940(1994).  
 RL Microbiology 140:2931-2940(1994).  
 CC -1- FUNCTION: CONVERTS LYSOPHOSPHATIDIC ACID (LPA) INTO PHOSPHATIDIC ACID BY INCORPORATING ACYL MOIETY AT THE 2 POSITION.  
 CC -1- CATALYTIC ACTIVITY: Acyl-CoA + 1-acyl-sn-glycerol 3-phosphate -> CoA + 1,2-diacyl-sn-glycerol 3-phosphate.  
 CC -1- PATHWAY: SECOND STEP IN DE NOVO PHOSPHOLIPID BIOSYNTHESIS.  
 CC -1- SIMILARITY: BELONGS TO THE 1-ACYL-SN-GLYCEROL-3-PHOSPHATE ACYLTRANSFERASE FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL: AE001117; AAC66417.1; -.  
 DR EMBL: L32861; AAC1407.1; -.  
 DR TIGR: B80037; -.  
 DR InterPro: IPR002123; Acyltransferase.  
 DR Pfam: PF01553; Acyltransferase; 1.  
 KM Phospholipid biosynthesis; transferase; Acyltransferase;  
 KM Complete proteome.  
 SQ SEQUENCE 250 AA; 28580 MW; A730BED7058AD999 CRC64;

Query Match 3.1%; Score 7; DB 1; Length 250;  
 Best Local Similarity 100.0%; Pred. No. 14;  
 Matches 7; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

OY 159 IPEPGR 165  
 DB 158 IPEPGR 164

RESULT 14

PSD8\_CAEL STANDARD; PRT: 250 AA.  
 ID PSD8\_CAEL  
 AC 023449;  
 DT 15-DEC-1998 (Rel. 37, Created)  
 DT 15-DEC-1998 (Rel. 37, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Probable 26S proteasome regulatory subunit S14.  
 GN ZK20.5.  
 OS Caenorhabditis elegans.  
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;  
 OC Rhabditidae; Peloderinae; Caenorhabditis.  
 OX NCBI\_TaxID=6239;  
 RN (1)  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=BRISTOL N2;  
 RA Gajadry S.;  
 RL Submitted (MAR-1996) to the EMBL/GenBank/DBJ databases.  
 CC -1- FUNCTION: ACTS AS A REGULATORY SUBUNIT OF THE 26 PROTEASOME WHICH IS INVOLVED IN THE ATP-DEPENDENT DEGRADATION OF UBIQUITINATED PROTEINS (BY SIMILARITY).  
 CC -1- SIMILARITY: BELONGS TO THE PROTEASOME SUBUNIT S14 FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL: Z69904; CA93778.1; -.  
 DR WormPep: ZK20.5; CE06608.  
 KW Proteasome.  
 SQ SEQUENCE 250 AA; 28795 MW; 7406538A0CE1B2DE CRC64;

Query Match 3.1%; Score 7; DB 1; Length 250;  
 Best Local Similarity 100.0%; Pred. No. 14;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 OY 171 TKVLAS 177  
 DB 30 TKVLAS 36

RESULT 15  
 PROC\_YEAST STANDARD; PRT: 286 AA.  
 ID PROC\_YEAST  
 AC P32263;  
 DT 01-OCT-1993 (Rel. 27, Created)  
 DT 01-OCT-1993 (Rel. 27, Last sequence update)  
 DT 15-JUL-1999 (Rel. 38, Last annotation update)  
 DE Pyruvate-5-carboxylate reductase (EC 1.5.1.2) (P5CR) (P5C reductase).  
 GN P303 OR ORE2 OR YER023W.  
 OS Saccharomyces cerevisiae (Baker's yeast).  
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;  
 OC Saccharomycetales; Saccharomycetaceae; Saccharomyces.  
 OX NCBI\_TaxID=4932;  
 RN (1)  
 RP SEQUENCE FROM N.A.  
 RC MEDLINE=92276362; PubMed=1592829;  
 RX MEDLINE=92276362; PubMed=1592829;  
 RA Brandiss M.C., Falvey D.A.;  
 RA "Proline biosynthesis in Saccharomyces cerevisiae: analysis of the RT P303 gene, which encodes delta 1-pyrroline-5-carboxylate reductase."  
 RL J. Bacteriol. 174:3782-3788(1992).  
 RN (2)  
 RP ERRATUM.  
 RX MEDLINE=92332464; PubMed=1352771;  
 RA Brandiss M.C., Falvey D.A.;  
 RA J. Bacteriol. 174:5176-5176(1992).  
 RN (3)  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=92374987; PubMed=1508147;

RA Neuville P., Aigle M.;  
 RT "ore2, a mutation affecting proline biosynthesis in the yeast  
 RT *Saccharomyces cerevisiae*, leads to a cdc phenotype.";   
 RL Mol. Genet. 234:193-200(1992).

RN [4]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-S288C / AB972;

RA Dietrich F.S., Mulligan J.T., Hennessey K.M., Allen E., Araujo R.,  
 RA Aviles E., Berno A., Brennan T., Carpenter J., Chen E., Cherry J.M.,  
 RA Chung E., Duncan M., Guzman E., Hartzell G., Hunnicke-Smith S.,  
 RA Hyman R., Kayser A., Komp C., Lashkari D., Lew H., Lin D.,  
 RA Mosedale D., Nakahara K., Namath A., Norgren R., Oefner P., Oh C.,  
 RA Petel F.X., Roberts D., Sehl P., Schramm S., Shogren T., Smith V.,  
 RA Taylor P., Wei Y., Yelton M., Botstein D., Davis R.W.;  
 RL Submitted (DEC-1994) to the EMBL/Genbank/DBJ databases.

CC -1 CATALYTIC ACTIVITY: L-proline + NAD(P)<sup>+</sup> = L-pyrroline-5-  
 CC cardoxylate + NAD(P)H.  
 CC -1 PATHWAY: TERMINAL (THIRD) STEP IN PROLINE BIOSYNTHESIS.

CC -1 SUBUNIT: HOMOTETRAMER.

CC -1 SIMILARITY: BELONGS TO THE PYRROLINE-5-CARBOXYLATE REDUCTASE  
 CC FAMILY.

CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).

CC -----  
 CC DR EMBL; M57886; AAA34905.1; -;  
 DR EMBL; X57338; CAA40614.1; -;  
 DR EMBL; U18778; AAB64556.1; -;  
 DR PIR; A41906; A41906.  
 DR PIR; A42722; A42722.  
 DR SCD; S000825; S25293.  
 DR PIR; S25293; S25293.  
 DR InterPro: IPR000304; PRO3.  
 DR Pfam: PF01089; P5CR.1.  
 DR PROSITE; PS00521; P5CR.1.  
 KM Oxidoreductase; Proline biosynthesis; NADP.  
 SQ SEQUENCE 286 AA; 30132 MW; AEB7ID93B46D08B3 CRC64;

Query Match 3 1%; Score 7; DB 1; Length 286;

Best Local Similarity 100.0%; Pred. No. 16;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7 SYVLGCT 13  
 |||||  
 DB 84 SYVLGCT 90

Search completed: August 28, 2002, 11:27:20  
 Job time: 513 sec



THIS PAGE BLANK (10/10)

Connecting via Winsock to STN

Welcome to STN International! Enter x:x

LOGINID:ssap1642cxy

PASSWORD:  
TERMINAL (ENTER 1, 2, 3, OR ?):2

\*\*\*\*\* Welcome to STN International \*\*\*\*\*

NEWS 1 Web Page URLs for STN Seminar Schedule - N. America  
NEWS 2 "Ask CAS" for self-help around the clock  
NEWS 3 BELLSTEIN, Reload and Implementation of a New Subject Area  
NEWS 4 ZDB will be removed from STN  
NEWS 5 US Patent Applications available in IFICDB, IFIPAT, and IFIUDB  
NEWS 6 Records from IP.com available in CAPLUS, HCAPLUS, and ZCAPLUS  
NEWS 7 BIOSIS Gene Names now available in TOXCENTER  
NEWS 8 Federal Research in Progress (FEDRIP) now available  
NEWS 9 New e-mail delivery for search results now available  
NEWS 10 MEDLINE Reload  
NEWS 11 PCTFUL has been reloaded  
NEWS 12 FOREG no longer contains STANDARDS file segment  
NEWS 13 USAN to be reloaded July 28, 2002;  
NEWS 14 saved answer sets no longer valid  
NEWS 15 Enhanced polymer searching in REGISTRY  
NEWS 16 NEFIRST to be removed from STN  
NEWS 17 CANCERLIT reload  
NEWS 18 PHARMAMarketletter (PHARMAML) - new on STN  
NEWS 19 NTIS has been reloaded and enhanced  
NEWS 20 JAPLO to be reloaded August 25, 2002  
NEWS 21 Aquatic Toxicity Information Retrieval (AQUIRE)  
NEWS 22 now available on STN  
NEWS 23 IFIPAT, IFICDB, and IFIUDB have been reloaded  
NEWS 24 The MEDLINE file segment of TOXCENTER has been reloaded  
NEWS 25 Sequence searching in REGISTRY enhanced  
NEWS EXPRESS February 1 CURRENT WINDOWS VERSION IS V6.0d,  
CURRENT MACINTOSH VERSION IS V6.0a(ENG) AND V6.0a(JP),  
AND CURRENT DISCOVER FILE IS DATED 05 FEBRUARY 2002  
NEWS HOURS STN Operating Hours Plus Help Desk Availability  
NEWS INTER General Internet Information  
NEWS LOGIN Welcome Banner and News Items  
NEWS PHONE Direct Dial and Telecommunication Network Access to STN  
NEWS WWW CAS World Wide Web Site (general information)

Enter NEWS followed by the item number or name to see news on that specific topic.

All use of STN is subject to the provisions of the STN Customer agreement. Please note that this agreement limits use to scientific research. Use for software development or design or implementation of commercial gateways or other similar uses is prohibited and may result in loss of user privileges and other penalties.

\*\*\*\*\* STN Columbus \*\*\*\*\*

FILE 'HOME' ENTERED AT 12:44:08 ON 28 AUG 2002

=> file medline, cancerlit, caplus, biosis, confsci, unpatfull, embase  
COST IN U.S. DOLLARS  
FULL ESTIMATED COST  
ENTRY 0.42  
TOTAL SESSION 0.42

FILE 'MEDLINE' ENTERED AT 12:45:03 ON 28 AUG 2002

FILE 'CANCERLIT' ENTERED AT 12:45:03 ON 28 AUG 2002

FILE 'CAPLUS' ENTERED AT 12:45:03 ON 28 AUG 2002  
USE IS SUBJECT TO THE TERMS OF YOUR STN CUSTOMER AGREEMENT.  
PLEASE SEE "HELP USAGETTERMS" FOR DETAILS.

COPYRIGHT (C) 2002 AMERICAN CHEMICAL SOCIETY (ACS)

FILE 'BIOSIS' ENTERED AT 12:45:03 ON 28 AUG 2002  
COPYRIGHT (C) 2002 BIOLOGICAL ABSTRACTS INC. (R)

FILE 'CONFSCI' ENTERED AT 12:45:03 ON 28 AUG 2002  
COPYRIGHT (C) 2002 Cambridge Scientific Abstracts (CSA)

FILE 'USPATFULL' ENTERED AT 12:45:03 ON 28 AUG 2002  
CA INDEXING COPYRIGHT (C) 2002 AMERICAN CHEMICAL SOCIETY (ACS)

FILE 'EMBASE' ENTERED AT 12:45:03 ON 28 AUG 2002  
COPYRIGHT (C) 2002 Elsevier Science B.V. All rights reserved.

=> s PGI  
L1 1251 PGI

=> s l1 and prostate?  
L2 28 l1 AND PROSTATE?

=> d 12 1-28

L2 ANSWER 1 OF 28 MEDLINE  
AN 2001400162 MEDLINE  
DN 21344590 Pubmed ID: 11443539  
TI Linkage and association studies of prostate cancer  
AU Xu J; Zheng S L; Hawkins G A; Faith D A; Kelly B; Isaacs S D; Wiley K E;  
Chang B; Ewing C M; Bujnowsky P; Carpen J D; Bleeker E R; Walsh P C;  
Trent J M; Meyers D A; Isaacs W B  
CS Center for Human Genomics, Wake Forest University School of Medicine,  
Winston-Salem, NC, USA.  
NC C458236 (NCT)  
SO AMERICAN JOURNAL OF HUMAN GENETICS, (2001 Aug) 69 (2) 341-50.  
CY Journal code: 0370475. ISSN: 0002-9297.  
DT United States  
LA Journal: Article; (JOURNAL ARTICLE)  
FS English  
OS Priority Journals  
OM OMIM-176807; OMIM-300147; OMIM-601518; OMIM-602759; OMIM-603688;  
EM OMIM-605367  
ED Entered STN: 20010820  
Last Updated on STN: 20010820  
Entered Medline: 20010816  
L2 ANSWER 2 OF 28 CANCERLIT  
AN 2002066870 CANCERLIT  
DN 21344590 Pubmed ID: 11443539  
TI Linkage and association studies of prostate cancer  
susceptibility: evidence for linkage at 8p22-21.

**THIS PAGE BLANK (USPTO)**

AU Xu J.; Zheng S. L.; Hawkins G. A.; Faich D. A.; Kelly B.; Isaacs S. D.; Wiley K. E.;  
 Chang B.; Ewing C. M.; Bujnovsky P.; Carpten J. D.; Bleecker E. R.; Walsh P. C.;  
 Trent J. M.; Meyers D. A.; Isaacs W. B.  
 Center for Human Genomics, Wake Forest University School of Medicine,  
 Winston-Salem, NC, USA.  
 NC CAS9236 (NCL)  
 SO JOURNAL OF HUMAN GENETICS, (2001 Aug) 69 (2) 341-50.  
 Journal code: 0370475. ISSN: 0002-9297.  
 CY United States  
 DT Journal; Article; (JOURNAL ARTICLE)  
 LA English  
 FS MEDLINE; Priority Journals  
 OS MEDLINE 2001400162; OMIM-176807; OMIM-300147; OMIM-601518; OMIM-602759;  
 OMIM-603688; OMIM-605367  
 EM 200108  
 ED Entered STM: 20020726  
 Last Updated on STM: 20020726  
 L2 ANSWER 3 OF 28 CAPLUS COPYRIGHT 2002 ACS  
 AN 2002:72326 CAPLUS  
 DN 136:129959  
 TI Human biallelic marker maps and their uses  
 IN Cohen, Daniel; Blumenfeld, Marta; Chumakov, Ilya; Abderrahim, Hadi;  
 PA Bihain, Bernard  
 Gensec, Fr.  
 SO PCT Int. Appl., 311 pp.  
 CODEN: PIXD2  
 DT Patent  
 LA English  
 FAN.CNT 1  
 PATENT NO. KIND DATE APPLICATION NO. DATE  
 WO 2002006525 A2 20020124 WO 2001-1B1477 20010628  
 W: AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN,  
 CR, CU, CZ, DE, DK, DM, DZ, EE, ES, FI, GB, GD, GE, GH, GM, HR,  
 HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT,  
 LU, LV, MA, MD, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM,  
 SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN,  
 YU, ZA, ZW, AM, AZ, BY, BG, KZ, MD, RU, TD, TG  
 RM: GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZW, AT, BE, CH, CY,  
 DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, TR, BF,  
 BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG  
 PRAI US 2000-219704P P 20000718  
 L2 ANSWER 4 OF 28 CAPLUS COPYRIGHT 2002 ACS  
 AN 2001:613738 CAPLUS  
 DN 136:197777  
 TI Linkage and association studies of prostate cancer  
 susceptibility: Evidence for linkage at 8p22-23  
 AU Xu, Jianfeng; Zheng, Sigun L.; Hawkins, Gregory A.; Faich, Dennis A.;  
 Kelly, Brian; Isaacs, Sarah D.; Wiley, Kathleen E.; Chang, Bao-Ji; Ewing,  
 Charles M.; Bujnovsky, Pirooska; Carpten, John D.; Bleecker, Eugene R.;  
 Walsh, Patrick C.; Trent, Jeffrey M.; Meyers, Deborah A.; Isaacs, William  
 B.  
 Center for Human Genomics, Wake Forest University School of Medicine,  
 Winston-Salem, NC, USA  
 SO AMERICAN JOURNAL OF HUMAN GENETICS (2001), 69(2), 341-350  
 CODEN: AJHGAG; ISSN: 0002-9297  
 PB University of Chicago Press  
 DT Journal  
 LA English  
 RE.CNT 36  
 ALL CITATIONS AVAILABLE IN THE RE FORMAT

L2 ANSWER 5 OF 28 CAPLUS COPYRIGHT 2002 ACS  
 AN 2001:537521 CAPLUS  
 DN 135:135934  
 TI The FGL gene associated with prostate cancer and  
 biallelic markers for assessment of prostate cancer risk  
 IN Cohen, Daniel; Blumenfeld, Marta; Chumakov, Ilya; Bouguetelret, Lydie  
 PA Gensec, Fr.  
 SO U.S., 366 pp., Cont.-in-part of U.S. Ser. No. 218,207.  
 CODEN: USXXAM  
 DT Patent  
 LA English  
 FAN.CNT 2  
 PATENT NO. KIND DATE APPLICATION NO. DATE  
 US 6265546 B1 20010724 US 1999-338907 19990623  
 US 5945522 A 19990831 US 1997-996306 19971222  
 EP 1052292 A1 20001115 EP 2000-115991 19981222  
 R: AT, BE, CH, DE, DK, ES, FR, GB, GR, IT, LI, LU, NL, SE, MC, PT,  
 IE, FI  
 PRAI US 1997-996306 A2 20020212 US 1998-218207 19981222  
 US 1998-99658P P 19980909  
 US 1998-218207 A2 19981222  
 EP 1998-960061 A3 19981222  
 RE.CNT 52  
 THERE ARE 52 CITED REFERENCES AVAILABLE FOR THIS RECORD  
 ALL CITATIONS AVAILABLE IN THE RE FORMAT  
 L2 ANSWER 6 OF 28 CAPLUS COPYRIGHT 2002 ACS  
 AN 2001:187442 CAPLUS  
 DN 134:276382  
 TI Toward a catalog of human genes and proteins: sequencing and analysis of  
 500 novel complete protein coding human cDNAs  
 AU Wiemann, Stefan; Weil, Bernd; Wellenreuther, Ruth; Gassenhuber, Johannes;  
 Glasel, Sabine; Ansorge, Wilhelm; Bocher, Michael; Blocker, Helmut;  
 Buerstch, Stefan; Blum, Helmut; Lauber, Jürgen; Nottmann, Andreas;  
 Beyer, Andreas; Köhler, Karl; Strack, Norbert; Mewes, Hans-Werner;  
 Oltewald, Birgit; Obermaier, Brigitte; Tampe, Jens; Heubner, Dagmar;  
 Wamulit, Rolf; Korn, Bernhard; Klein, Michaela; Poustka, Annemarie  
 Molecular Genome Analysis, German Cancer Research Center, Heidelberg,  
 69120, Germany  
 SO Genome Research (2001), 11(3), 422-435  
 CODEN: GREERS; ISSN: 1088-9051  
 PB Cold Spring Harbor Laboratory Press  
 DT Journal  
 LA English  
 RE.CNT 51  
 THERE ARE 51 CITED REFERENCES AVAILABLE FOR THIS RECORD  
 ALL CITATIONS AVAILABLE IN THE RE FORMAT  
 L2 ANSWER 7 OF 28 CAPLUS COPYRIGHT 2002 ACS  
 AN 1999:421796 CAPLUS  
 DN 131:83975  
 TI Cloning, expression and characterization of prostate cancer gene  
 (PGL) and use in prostate cancer diagnosis  
 IN Cohen, Daniel; Blumenfeld, Marta; Chumakov, Ilya; Bouguetelret, Lydie  
 PA Gensec, Fr.  
 SO PCT Int. Appl., 385 pp.  
 CODEN: PIXD2  
 DT Patent  
 LA English  
 FAN.CNT 2  
 PATENT NO. KIND DATE APPLICATION NO. DATE  
 WO 9932644 A2 19990701 WO 1998-1B2133 19981222  
 WO 9932644 A3 19990910

**THIS PAGE BLANK (USPTO)**

W: AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CU, CZ, DE, DE, EE, ES, FI, GB, GD, GE, GH, GM, GR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LV, MD, MG, MK, MN, MM, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, ST, TM, TR, TT, UA, UG, US, UZ, VN, YU, ZW, AM, AZ, BY, BG, KZ, MD, RU, TJ, TM, RW, GH, GM, KE, LS, MW, SD, SZ, UG, ZW, AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG

US 5945522 A 19990831 US 1997-966306 19971222  
CA 2311416 A 19990701 CA 1998-2311416 19981222  
AU 9915740 A 19990701 AU 1998-15740 19981222  
EP 991770 A 20000412 EP 1998-966061 19981222  
R: AT, BE, CH, DE, DK, ES, FR, GB, GR, IT, LI, LU, NL, SE, MC, PT, IE, FI

EP 1052292 A1 20001115 EP 2000-115991 19981222  
R: AT, BE, CH, DE, DK, ES, FR, GB, GR, IT, LI, LU, NL, SE, MC, PT, IE, FI

AT 197817 E 20001215 AT 1998-966061 19981222  
JP 2002516657 T2 20020611 JP 2000-525562 19981222  
US 1997-966306 A 19971222  
US 1998-96658P P 19980909  
WO 1998-182133 W 19981222

L2 ANSWER 8 OF 28 CAPLUS COPYRIGHT 2002 ACS  
AN 1999-81619 CAPLUS  
DN 130119526  
TI Biallelic markers for use in constructing a high density disequilibrium map of the human genome  
IN Cohen, Daniel; Blumenfeld, Marta; Tchoumakov, Ilya  
PA Genset, Fr.  
SO PCT Int. Appl., 288 pp.  
CODEN: PIXXD2  
DT Patent  
LA English  
FAM, CNT 3

PATENT NO.	KIND	DATE	APPLICATION NO.	DATE
WO 9904038	A2	19990128	WO 1998-181193	19980717
WO 9904038	A3	19990610		
W: AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CU, CZ, DE, DK, EE, ES, FI, GB, GR, GM, GR, HU, ID, IL, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MM, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, ST, TM, TR, TT, UA, UG, US, UZ, VN, YU, ZW, AM, AZ, BY, BG, KZ, MD, RU, TJ, TM, RW, GH, GM, KE, LS, MW, SD, SZ, UG, ZW, AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG				
AU 9884569	A1	19990210	AU 1998-84569	19980717
AU 746682	B2	20020502		
EP 1002131	A2	20000524	EP 1998-935225	19980717
R: AT, BE, CH, DE, DK, ES, FR, GB, GR, IT, LI, LU, NL, SE, MC, PT, IE, FI				
PRAI US 1998-82614P	P	19980421		
EP 1997-401740	A	19970718		
WO 1998-181193	W	19980717		

L2 ANSWER 9 OF 28 CAPLUS COPYRIGHT 2002 ACS  
AN 1993-140975 CAPLUS  
DN 118-140975  
TI Rat gene mapping using PCR-analyzed microsatellites  
AU Serikawa, Tadao; Kuramoto, Takashi; Hilbert, Pascale; Mori, Masayuki;

Yamada, Junzo; Dubay, Christopher J.; Lindpaintner, Klaus; Ganten, Detlev; Genset, Jean Louis; et al  
Proc. Med. Kyoto Univ., Kyoto, 606, Japan  
Genetics (1992), 131(3), 701-21  
CODEN: GENTAE; ISSN: 0016-6731

DT Journal  
LA English

L2 ANSWER 10 OF 28 BIOSIS COPYRIGHT 2002 BIOLOGICAL ABSTRACTS INC.  
AN 2002-194891 BIOSIS  
DN PREV200200194891  
TI Prostate cancer gene  
AU Cohen, Daniel (1); Blumenfeld, Marta; Chumakov, Ilya; Bougueleret, Lydie  
CS (1) Fontenay-sous-bois France  
ASSIGNEE: Genset, France  
PI US 6346381 February 12, 2002  
SO Official Gazette of the United States Patent and Trademark Office Patents, (Feb. 12, 2002) Vol. 125, No. 2, pp. No Pagination.  
http://www.uspto.gov/web/menu/patdata.html. e-file.  
ISSN: 0098-1133.  
DT Patent  
LA English

L2 ANSWER 11 OF 28 BIOSIS COPYRIGHT 2002 BIOLOGICAL ABSTRACTS INC.  
AN 2001-432446 BIOSIS  
DN PREV200100432446  
TI Prostate cancer gene.  
AU Cohen, Daniel (1); Blumenfeld, Marta; Chumakov, Ilya; Bougueleret, Lydie  
CS (1) Neuilly sur Seine France  
ASSIGNEE: Genset, France  
PI US 6265546 July 24, 2001  
SO Official Gazette of the United States Patent and Trademark Office Patents, (July 24, 2001) Vol. 1248, No. 4, pp. No Pagination. e-file.  
ISSN: 0098-1133.  
DT Patent  
LA English

L2 ANSWER 12 OF 28 BIOSIS COPYRIGHT 2002 BIOLOGICAL ABSTRACTS INC.  
AN 2001-395125 BIOSIS  
DN PREV200100395125  
TI Linkage and association studies of prostate cancer susceptibility: Evidence for linkage at 8p22-23.  
AU Xu, Jianteng; Zheng, Siqun L.; Hawkins, Gregory A.; Faith, Dennis A.; Kelly, Brian; Isaacs, Sarah D.; Wiley, Kathleen E.; Chang, Bao-li; Swing, Charles W.; Bujnovszky, Pirooska; Carpen, John D.; Blecker, Eugene R.; Walsh, Patrick C.; Trent, Jeffrey M.; Meyers, Deborah A.; Isaacs, William B. (1)  
CS (1) Johns Hopkins Hospital, 600 N. Wolfe Street, Marburg 115, Baltimore, MD, 21287; wisacs@hmi.edu USA  
SO American Journal of Human Genetics, (August, 2001) Vol. 69, No. 2, pp. 341-350. print.  
ISSN: 0002-9297.  
DT Article  
LA English  
SL English

L2 ANSWER 13 OF 28 BIOSIS COPYRIGHT 2002 BIOLOGICAL ABSTRACTS INC.  
AN 1999-1525274 BIOSIS  
DN PREV199900525274  
TI Prostate cancer gene.  
AU Cohen, Daniel (1); Chumakov, Ilya; Blumenfeld, Marta; Bougueleret, Lydie  
CS (1) Fontenay-sous-bois France  
ASSIGNEE: GENSET  
PI US 5945522 Aug. 31, 1999

**THIS PAGE BLANK (USPTO)**



SO Official Gazette of the United States Patent and Trademark Office Patents,  
(Aug. 31, 1999) Vol. 1225, No. 5, pp. NO PAGINATION.  
ISSN: 0098-1133.

DT Patent  
LA English  
L2 ANSWER 14 OF 28 USPTAFULL  
AN 2002:157048 USPTAFULL  
TI APOPTOSIS INDUCING MOLECULE II AND METHODS OF USE  
IN EBNER, REINHARD, GAITHERSBURG, MD, UNITED STATES  
YU, GUO-LIANG, BERKELEY, CA, UNITED STATES  
RUBEN, STEVEN M., OLNEY, MD, UNITED STATES  
ZHANG, JUN, BETHESDA, MD, UNITED STATES  
ULBRICH, STEPHEN, ROCKVILLE, MD, UNITED STATES  
ZHAH, YIFAN, GAITHERSBURG, MD, UNITED STATES  
PA Human Genome Sciences (U.S. corporation)  
PI US 2002061647 AI 20020627  
AI US 1999-252656  
RLI Continuation-in-part of Ser. No. US 1998-27287, filed on 20 Feb 1998,  
PENDING Continuation-in-part of Ser. No. US 1998-3886, filed on 7 Jan  
1998, ABANDONED Continuation-in-part of Ser. No. US 1997-822953, filed  
on 21 Mar 1997, ABANDONED  
PRAI US 1998-75409P 19980220 (60)  
US 1996-13923P 19960322 (60)  
US 1996-30157P 19961031 (60)  
DT Utility  
FS APPLICATION  
LN.CNT 6195  
INCL INCLM: 435/069.100  
NCL INCLM: 530/350.000; 530/399.000; 514/012.000; 536/023.500  
IC NCLM: 435/069.100  
NCLM: 530/350.000; 530/399.000; 514/012.000; 536/023.500  
[7]  
ICM: A61K038-18  
ICS: C12P021-06; C07H021-04; C07K014-00; G01N033-53  
CAS INDEXING IS AVAILABLE FOR THIS PATENT.  
L2 ANSWER 15 OF 28 USPTAFULL  
AN 2002:126357 USPTAFULL  
TI APOPTOSIS INDUCING MOLECULE II  
IN EBNER, REINHARD, GAITHERSBURG, MD, UNITED STATES  
YU, GUO-LIANG, DARNESTOWN, MD, UNITED STATES  
RUBEN, STEVEN M., OLNEY, MD, UNITED STATES  
ULBRICH, STEPHEN, ROCKVILLE, MD, UNITED STATES  
PA Human Genome Sciences, Inc. (U.S. corporation)  
PI US 2002064869 AI 20020530  
AI US 1998-27287 AI 19980220 (9)  
RLI Continuation-in-part of Ser. No. US 1997-822953, filed on 21 Mar 1997,  
ABANDONED  
PRAI US 1996-30157P 19961031 (60)  
US 1996-13923P 19960322 (60)  
DT Utility  
FS APPLICATION  
LN.CNT 4242  
INCL INCLM: 435/320.100  
NCL INCLM: 435/069.100; 435/325.000; 536/023.500  
NCLM: 435/320.100  
NCLM: 435/069.100; 435/325.000; 536/023.500  
[7]  
ICM: C12N015-63  
ICS: C07H021-04; C12N015-00; C12N015-74; C12N005-06; C12N015-70;  
C12N015-09  
CAS INDEXING IS AVAILABLE FOR THIS PATENT.

L2 ANSWER 16 OF 28 USPTAFULL  
AN 2002:61244 USPTAFULL  
TI Mutated cyclin G1 protein  
IN Gordon, Erlinda Maria, Glendale, CA, UNITED STATES  
HALL, Frederick L., Glendale, CA, UNITED STATES  
PI US 2002035079 AI 20020321  
AI US 2001-796149 AI 20010228 (9)  
RLI Continuation-in-part of Ser. No. US 2000-325765, filed on 2 Mar 2000,  
ABANDONED  
DT Utility  
FS APPLICATION  
LN.CNT 1822  
INCL INCLM: 514/044.000  
NCL INCLM: 424/094.100  
NCLM: 514/044.000  
NCLM: 424/094.100  
[7]  
ICM: A61K048-00  
ICS: A61K038-43  
CAS INDEXING IS AVAILABLE FOR THIS PATENT.  
L2 ANSWER 17 OF 28 USPTAFULL  
AN 2002:29243 USPTAFULL  
TI Prostate cancer gene  
IN Cohen, Daniel, Fontenay-sous-bois, FRANCE  
Blumenfeld, Marita, Paris, FRANCE  
Chumakov, Ilya, Vaux-le-Penil, FRANCE  
Bougueleret, Lydie, Vanves, FRANCE  
PA Geneset, FRANCE (non-U.S. corporation)  
PI US 6346381 BI 19981222 (9)  
AI US 1998-218207  
RLI Continuation-in-part of Ser. No. US 1997-996306, filed on 22 Dec 1997,  
now patented Pat. No. US 5945522  
PRAI US 1998-99658P 19980909 (60)  
DT Utility  
FS GRANTED  
LN.CNT 17612  
INCL INCLM: 435/006.000  
NCL INCLM: 435/091.100; 435/091.200; 536/023.100; 536/024.300  
NCLM: 435/006.000  
NCLM: 435/091.100; 435/091.200; 536/023.100; 536/024.300  
[7]  
ICM: C120001-68  
ICS: C12P019-34; C07M021-02  
EXF 435/6; 435/91.2; 435/91.1; 536/23.1; 536/24.3  
CAS INDEXING IS AVAILABLE FOR THIS PATENT.  
L2 ANSWER 18 OF 28 USPTAFULL  
AN 2001:168248 USPTAFULL  
TI MYC homology region II--associated protein and uses thereof  
IN DePinho, Ronald A., Brookline, MA, United States  
PA Albert Einstein College of Medicine of Yeshiva University, Bronx, NY,  
United States (U.S. corporation)  
PI US 6297368 BI 20011002  
AI US 2000-497779 BI 20000203 (9)  
RLI Continuation-in-part of Ser. No. US 1997-946692, filed on 8 Oct 1997,  
now patented, Pat. No. US 6040425  
DT Utility  
FS GRANTED  
LN.CNT 514  
INCL INCLM: 536/023.500  
NCL INCLM: 530/350.000; 530/300.000; 530/827.000; 514/012.000  
NCLM: 536/023.500  
NCLM: 530/300.000; 530/350.000; 530/827.000

**THIS PAGE BLANK (USPTO)**

IC [7]  
 ICM: C07H021-04  
 ICS: C07K014-47  
 EXP 536/23.1; 536/23.5; 530/350  
 CAS INDEXING IS AVAILABLE FOR THIS PATENT.

L2 ANSWER 19 OF 28 USPATFULT  
 AN 2001:158014 USPATFULT  
 TI Methode, software and apparatus for identifying genomic regions harboring a gene associated with a detectable trait  
 IN Schork, Nicholas J., Shaker Heights, OH, United States  
 Essoux, Laurent, Paris, France  
 Cohen-Akenine, Annick, Paris, France  
 Blumenfeld, Marta, Paris, France  
 Cohen, Daniel, Neuilly-sur-Seine, France  
 Gense, Paris, France (non-U.S. corporation)  
 PI US 6291182 BI 20010918  
 AI US 1999-438016 19991110 (9)  
 PRAI US 1998-107986P 19981110 (60)  
 US 1999-140785P 19990623 (60)  
 DT Utility  
 FS GRANTED  
 LN.CNT 4241  
 INCL INCLM: 435/006.000  
 INCLS: 702/027.000  
 NCL NCLM: 435/006.000  
 NCLS: 702/027.000  
 IC [7]  
 ICM: C120001-68  
 ICS: G01N031-00  
 EXP 435/6; 530/300; 530/324; 530/325; 530/326; 712/200; 702/27  
 CAS INDEXING IS AVAILABLE FOR THIS PATENT.

L2 ANSWER 20 OF 28 USPATFULT  
 AN 2001:117151 USPATFULT  
 TI Prostate cancer gene  
 IN Cohen, Daniel, Neuilly sur Seine, France  
 Blumenfeld, Marta, Paris, France  
 Chumakov, Ilya, Vaux-le-Penil, France  
 Bouguetere, Lydie, Vanves, France  
 Gense, France (non-U.S. corporation)  
 PI US 6265546 BI 20010724  
 AI US 1999-338907 19990623 (9)  
 RLI Continuation-in-part of Ser. No. US 1998-218207, filed on 22 Dec 1998, now patented, Pat. No. US 5945522  
 PRAI US 1998-99658P 19980909 (60)  
 US 1998-107986P 19981110 (60)  
 DT Utility  
 FS GRANTED  
 LN.CNT 7782  
 INCL INCLM: 530/350.000  
 INCLS: 424/174.100; 435/007.100; 514/002.000  
 NCL NCLM: 530/350.000  
 NCLS: 424/174.100; 435/007.100  
 IC [7]  
 ICM: C07K001-00  
 ICS: A61K039-395; A01N037-18; G01N035-53  
 EXP 530/350; 530/300; 530/324; 530/325; 514/2; 514/10; 514/12; 424/174.1; 435/7.1  
 CAS INDEXING IS AVAILABLE FOR THIS PATENT.

L2 ANSWER 21 OF 28 USPATFULT  
 AN 2001:86218 USPATFULT

TI Sequence specific DNA binding p53  
 IN Vogelstein, Bert, Baltimore, MD, United States  
 Kinzler, Kenneth W., Belair, MD, United States  
 Sherman, Michael I., Allendale, NJ, United States  
 The Johns Hopkins University, Baltimore, MD, United States (U.S. corporation)  
 PA Pharmagenics, Inc., Allendale, NJ, United States (U.S. corporation)  
 PI US 6245515 BI 20010612  
 AI US 1999-399773 19990921 (9)  
 RLI Division of Ser. No. US 1994-299074, filed on 1 Sep 1994, now patented, Pat. No. US 5955263 Division of Ser. No. US 1997-860758, filed on 31 Mar 1992, now patented, Pat. No. US 5862623 Continuation-in-part of Ser. No. US 1991-715182, filed on 14 Jun 1991, now abandoned  
 DT Utility  
 FS GRANTED  
 LN.CNT 1804  
 INCL INCLM: 435/006.000  
 INCLS: 536/023.100; 536/024.330; 536/024.500; 536/025.300; 536/027.000  
 NCL NCLM: 435/006.000  
 NCLS: 536/023.100; 536/024.330; 536/024.500; 536/025.300  
 IC [7]  
 ICM: C120001-68  
 ICS: C07H021-02; C07H021-04; C07H019-00  
 EXP 435/6; 536/23.1; 536/24.33; 536/24.5; 536/25.3; 536/27  
 CAS INDEXING IS AVAILABLE FOR THIS PATENT.

L2 ANSWER 22 OF 28 USPATFULT  
 AN 2000:117520 USPATFULT  
 TI G-protein coupled receptor protein and a DNA encoding the receptor  
 IN Hinuma, Shuji, Tsukuba, Japan  
 Hosoya, Masaki, Tsukuba, Japan  
 Fujii, Ryo, Tsukuba, Japan  
 Ohtaki, Tetsuya, Tsukuba, Japan  
 Fukusumi, Shoji, Tsukuba, Japan  
 Ongi, Kazuniro, Tsukuba, Japan  
 Takeda Chemical Industries, Ltd., Osaka, Japan (non-U.S. corporation)  
 PI US 6114139 20000905  
 WO 9605302 19960222  
 AI WO 1995-513974 19950914 (8)  
 WO 1995-JP1599 19950810  
 PRAI JP 1994-189272 19940811  
 JP 1994-189273 19940811  
 JP 1994-189274 19940811  
 JP 1994-236356 19940930  
 JP 1994-236357 19940930  
 JP 1994-270017 19941102  
 JP 1994-326611 19941228  
 JP 1995-7177 19950120  
 JP 1995-57186 19950316  
 JP 1995-93989 19950419  
 DT Utility  
 FS GRANTED  
 LN.CNT 1308  
 INCL INCLM: 435/069.100  
 INCLS: 530/350.000; 536/023.500; 435/325.000; 435/320.100  
 NCL NCLM: 435/069.100  
 NCLS: 435/320.100; 435/325.000; 530/350.000; 536/023.500  
 IC [7]  
 ICM: C12N015-12  
 ICS: C12N015-63; C12N005-10; C07K014-705  
 EXP 536/23.5; 435/69.1; 435/325; 435/320.1; 530/350  
 CAS INDEXING IS AVAILABLE FOR THIS PATENT.

**THIS PAGE BLANK (USPTO)**

L2 ANSWER 23 OF 28 USPTFUL  
AN 2000:34669 USPTFUL  
TI Myc homology region II--associated protein  
IN Depinho, Ronald A., Pelham Manor, NY, United States  
PA Albert Einstein College of Medicine of Yeshiva University, Bronx, NY,  
United States (U.S. corporation)  
PI US 6040425 20000321  
AI US 1997-946692 19971008 (8)  
DT Utility  
FS Granted  
LN.CNT 554  
INCL INCLM: 530/350.000  
NCLM: 530/300.000; 530/827.000; 514/012.000  
NCLM: 530/350.000  
NCLM: 530/300.000; 530/827.000  
IC ICM: C07K014-47  
ICS: A61K038-17  
EXF 530/300; 530/350; 530/827; 514/12  
CAS INDEXING IS AVAILABLE FOR THIS PATENT.

L2 ANSWER 24 OF 28 USPTFUL  
AN 1999:113551 USPTFUL  
TI Sequence specific DNA binding by p53  
IN Vogelstein, Bert, Baltimore, MD, United States  
Kinzler, Kenneth W., Baltimore, MD, United States  
Sherman, Michael I., Glen Ridge, NJ, United States  
Johns Hopkins University, Baltimore, MD, United States (U.S. corporation)  
PI US 5955263 19990921  
AI US 1994-299074 19940901 (8)  
RUI Division of Ser. No. US 1997-860758, filed on 31 Mar 1992, now patented, Pat. No. US 5362623 which is a continuation-in-part of Ser. No. US 1991-715182, filed on 14 Jun 1991, now abandoned  
DT Utility  
FS Granted  
LN.CNT 2012  
INCL INCLM: 435/006.000  
NCLM: 435/007.230; 436/063.000; 436/064.000; 536/023.100; 536/024.100  
NCLM: 435/006.000  
NCLM: 435/007.230; 436/063.000; 436/064.000; 536/023.100; 536/024.100  
IC ICM: C120001-68  
ICS: G01N033-574; G01N033-48; C07H021-02  
EXF 435/6; 435/7.23; 435/7.92; 536/23.5; 536/23.1; 536/24.1; 436/63; 436/64  
CAS INDEXING IS AVAILABLE FOR THIS PATENT.

L2 ANSWER 25 OF 28 USPTFUL  
AN 1999:102911 USPTFUL  
TI Prostate cancer gene  
IN Cohen, Daniel, Fontenay-sous-bois, France  
Chumakov, Ilya, Vaux-le-Penil, France  
Blumenfeld, Marta, Paris, France  
Bougueleret, Lydie, Vanves, France  
PI GENSET, Paris, France (non-U.S. corporation)  
AI US 5945552 19990831  
DT US 1997-996306 19971222 (8)  
FS Utility  
LN.CNT 5177  
INCL INCLM: 536/023.100  
NCLM: 435/006.000; 536/024.100; 536/024.310; 536/024.320; 536/024.330;

NCLM: 435/006.000; 536/024.100; 536/024.300; 536/024.310; 536/024.320;  
NCLM: 435/006.000; 536/024.330  
IC ICM: C07H021-02  
ICS: C07H021-04; C120001-68  
EXF 435/6; 536/24.3; 536/24.31; 536/24.32; 536/24.33; 536/24.1; 536/23.1  
CAS INDEXING IS AVAILABLE FOR THIS PATENT.

L2 ANSWER 26 OF 28 USPTFUL  
AN 97:51869 USPTFUL  
TI Isolated nucleic acid encoding a ubiquitous nuclear receptor  
IN Liao, Shursung, Chicago, IL, United States  
Song, Ching, Durham, NC, United States  
PA Arch Development Corporation, Chicago, IL, United States (U.S. corporation)  
PI US 5639616 19970617  
AI US 1994-342411 19941118 (8)  
RUI Continuation-in-part of Ser. No. US 1993-152003, filed on 10 Nov 1993, now abandoned  
DT Utility  
FS Granted  
LN.CNT 4472  
INCL INCLM: 435/007.100  
NCLM: 435/069.100; 435/252.300; 435/320.100; 536/023.500; 536/024.300  
NCLM: 435/007.100  
NCLM: 435/069.100; 435/252.300; 435/320.100; 536/023.500; 536/024.300  
IC ICM: C12N005-10  
ICS: C12N015-12  
EXF 435/67.1; 435/69.1; 435/252.3; 435/320.1; 536/23.5; 536/24.3  
CAS INDEXING IS AVAILABLE FOR THIS PATENT.

L2 ANSWER 27 OF 28 USPTFUL  
AN 94:97462 USPTFUL  
TI Sequence specific DNA binding by p53  
IN Vogelstein, Bert, Baltimore, MD, United States  
Kinzler, Kenneth W., Baltimore, MD, United States  
Sherman, Michael I., Glen Ridge, NJ, United States  
The John Hopkins University, Baltimore, MD, United States (U.S. corporation)  
PI US 5362623 19941108  
AI US 1992-860758 19920331 (7)  
RUI Continuation-in-part of Ser. No. US 1991-715182, filed on 14 Jun 1991  
DT Utility  
FS Granted  
LN.CNT 1741  
INCL INCLM: 435/006.000  
NCLM: 435/320.100; 536/024.100; 536/024.310; 424/002.000  
NCLM: 435/006.000  
NCLM: 435/320.100; 536/024.100; 536/024.310  
IC ICM: C120001-68  
ICS: 435/6; 435/7.1; 435/7.21; 435/7.23; 435/7.92; 435/7.94; 435/70.3; 435/272; 435/975; 435/320.1; 436/501; 436/518; 436/63; 436/64; 436/811; 436/813; 424/2; 536/23.1; 536/24.31; 536/24.1  
EXF 435/6; 435/7.1; 435/7.21; 435/7.23; 435/7.92; 435/7.94; 435/70.3; 435/272; 435/975; 435/320.1; 436/501; 436/518; 436/63; 436/64; 436/811; 436/813; 424/2; 536/23.1; 536/24.31; 536/24.1  
CAS INDEXING IS AVAILABLE FOR THIS PATENT.

L2 ANSWER 28 OF 28 EMBASE COPYRIGHT 2002 ELSEVIER SCI. B.V.  
AN 2001267256 EMBASE  
TI Linkage and association studies of prostate cancer susceptibility. Evidence for linkage at 8p22-23.

**THIS PAGE BLANK (USPTO)**

AU Xu J.; Zheng S.L.; Hawkins G.A.; Faith D.A.; Kelly B.; Isaacs S.D.; Willey K.E.; Chang B.-L.; Ewing C.M.; Bujnovszky P.; Carpten J.D.; Bleecker E.R.; Walsh P.C.; Trent J.M.; Meyers D.A.; Isaacs W.B.  
 CS Dr. W.B. Isaacs, Marburg 115, Johns Hopkins Hospital, 600 N. Wolfe Street, Baltimore, MD 21287, United States. wisacs@jhm1.edu  
 SO American Journal of Human Genetics, (2001) 69/2 (341-350).  
 Refs: 36  
 ISSN: 0002-9297 CODEN: AJHGAG  
 CY United States  
 DT Journal; Article  
 FS 016 Cancer  
 022 Human Genetics  
 028 Urology and Nephrology  
 LA English  
 SL English  
  
 => d 12 1-28d his  
 '1-28d' IS NOT VALID HERE  
 For an explanation, enter "HELP DISPLAY HISTORY".  
 => d his  
 (FILE 'HOME' ENTERED AT 12:44:08 ON 28 AUG 2002)  
 FILE 'MEDLINE, CANCERLIT, CAPLUS, BIOSIS, COMPSCT, USPATFULL, EMBASE'  
 ENTERED AT 12:45:03 ON 28 AUG 2002  
 L1 1251 S PGI  
 L2 28 S L1 AND PROSTATE?  
 L3 => s antibod? and 11  
 120 ANTIBOD? AND L1  
 L4 => s 13 and prostate?  
 16 L3 AND PROSTATE?  
 L5 => duprem 14  
 DUPREM IS NOT A RECOGNIZED COMMAND  
 The previous command name entered was not recognized by the system.  
 For a list of commands available to you in the current file, enter  
 "HELP COMMANDS" at an arrow prompt (=>).  
 => dup rem 14  
 PROCESSING COMPLETED FOR L4  
 L5 16 DUP REM L4 (0 DUPLICATES REMOVED)  
 => d 15 1-16  
 L5 ANSWER 1 OF 16 USPATFULL  
 AN 2002:157048 USPATFULL  
 TI APOPTOSIS INDUCING MOLECULE II AND METHODS OF USE  
 IN EBNER, REINHARD, GAITHERSBURG, MD, UNITED STATES  
 YU, GUO-LIANG, BERKELEY, CA, UNITED STATES  
 RUBEN, STEVEN M., OLNEY, MD, UNITED STATES  
 ZHANG, JUN, BETHESDA, MD, UNITED STATES  
 ULBRICH, STEPHEN, ROCKVILLE, MD, UNITED STATES  
 ZHAI, YIFAN, GAITHERSBURG, MD, UNITED STATES  
 Human Genome Sciences (U.S. corporation)  
 PI US 2002081647 AI 20020627  
 P1 US 1999-252656 AI 19990219 (9)  
 R11 Continuation-in-part of Ser. No. US 1998-27287, filed on 20 Feb 1998,  
 PENDING Continuation-in-part of Ser. No. US 1998-3886, filed on 7 Jan  
 1998, ABANDONED Continuation-in-part of Ser. No. US 1997-822953, filed  
 on 21 Mar 1997, ABANDONED

PRA1 US 1998-75403P 19980220 (60)  
 US 1996-13923P 19960322 (60)  
 DT US 1996-30157P 19961031 (60)  
 FS Utility  
 IN APPLICATION  
 LN CNT 6195  
 INCL INCLM: 435/069.100  
 INCL INCLM: 530/350.000; 530/399.000; 514/012.000; 536/023.500  
 NCL NCLM: 435/069.100  
 NCLM: 530/350.000; 530/399.000; 514/012.000; 536/023.500  
 IC [7]  
 ICM: A61K038-18  
 ICS: C12P021-06; C07H021-04; C07K014-00; G01N033-53  
 CAS INDEXING IS AVAILABLE FOR THIS PATENT.  
 L5 ANSWER 2 OF 16 USPATFULL  
 AN 2002:126357 USPATFULL  
 TI APOPTOSIS INDUCING MOLECULE II  
 IN EBNER, REINHARD, GAITHERSBURG, MD, UNITED STATES  
 YU, GUO-LIANG, DARNESTOWN, MD, UNITED STATES  
 RUBEN, STEVEN M., OLNEY, MD, UNITED STATES  
 ULBRICH, STEPHEN, ROCKVILLE, MD, UNITED STATES  
 Human Genome Sciences, Inc. (U.S. corporation)  
 PI US 2002064869 AI 20020530  
 P1 US 1998-27287 AI 19980220 (9)  
 R11 Continuation-in-part of Ser. No. US 1997-822953, filed on 21 Mar 1997,  
 ABANDONED  
 PRA1 US 1996-30157P 19961031 (60)  
 US 1996-13923P 19960322 (60)  
 DT Utility  
 IN APPLICATION  
 LN CNT 4242  
 INCL INCLM: 435/320.100  
 INCL INCLM: 435/069.100; 435/325.000; 536/023.500  
 NCL NCLM: 435/320.100  
 NCLM: 435/069.100; 435/325.000; 536/023.500  
 IC [7]  
 ICM: C12N015-63  
 ICS: C07H021-04; C12N015-00; C12N015-74; C12N005-06; C12N015-70;  
 C12N015-09  
 CAS INDEXING IS AVAILABLE FOR THIS PATENT.  
 L5 ANSWER 3 OF 16 USPATFULL  
 AN 2002:61244 USPATFULL  
 TI Mutated cyclin G1 protein  
 IN Gordon, Erlinda Maria, Glendale, CA, UNITED STATES  
 Hall, Frederick L., Glendale, CA, UNITED STATES  
 PI US 2002035079 AI 20020321  
 P1 US 2001-796149 AI 20010228 (9)  
 R11 Continuation-in-part of Ser. No. US 2000-325765, filed on 2 Mar 2000,  
 ABANDONED  
 DT Utility  
 IN APPLICATION  
 LN CNT 1822  
 INCL INCLM: 514/044.000  
 INCL INCLM: 424/094.100  
 NCL NCLM: 514/044.000  
 NCLM: 424/094.100  
 IC [7]  
 ICM: A61K048-00  
 ICS: A61K038-43  
 CAS INDEXING IS AVAILABLE FOR THIS PATENT.  
 L5 ANSWER 4 OF 16 USPATFULL



**THIS PAGE BLANK (USPTO)**

AN 2002:29243 USPTFLL  
 TI Prostate cancer gene  
 IN Cohen, Daniel, Fontenay-sous-bois, FRANCE  
 Blumenfeld, Marta, Paris, FRANCE  
 Chumakov, Ilya, Vaux-le-Penil, FRANCE  
 Bougueleret, Lydie, Vanves, FRANCE  
 PA Genset, FRANCE (non-U.S. corporation)  
 PI US 6346381 B1 20020212  
 AI US 1998-218207 19981222 (9)  
 RLI Continuation-in-part of Ser. No. US 1997-996306, filed on 22 Dec 1997,  
 now patented, Pat. No. US 5945522  
 PRAI US 1998-99658P 19980909 (60)  
 DT Utility  
 FS GRANTED  
 LN CNT 17612  
 INCL INCLM: 435/006.000  
 INCLM: 435/091.100; 435/091.200; 536/023.100; 536/024.300  
 NCLM: 435/006.000  
 NCLM: 435/091.100; 435/091.200; 536/023.100; 536/024.300  
 IC [7]  
 ICM: C12Q001-68  
 EXF ICS: C12P019-34; C07M021-02  
 435/6; 435/91.2; 435/91.1; 536/23.1; 536/24.3  
 CAS INDEXING IS AVAILABLE FOR THIS PATENT.  
 L5 ANSWER 5 OF 16 USPTFLL  
 AN 2001:168248 USPTFLL  
 TI Myc homology region II--associated protein and uses thereof  
 IN Depinho, Ronald A., Brookline, MA, United States  
 PA Albert Einstein College of Medicine of Yeshiva University, Bronx, NY,  
 United States (U.S. corporation)  
 PI US 6297368 B1 20011002  
 AI US 2000-497779 20000203 (9)  
 RLI Continuation-in-part of Ser. No. US 1997-946692, filed on 8 Oct 1997,  
 now patented, Pat. No. US 6040425  
 DT Utility  
 FS GRANTED  
 LN CNT 514  
 INCL INCLM: 536/023.500  
 INCLM: 530/350.000; 530/300.000; 530/827.000; 514/012.000  
 NCLM: 536/023.500  
 NCLM: 530/300.000; 530/350.000; 530/827.000  
 IC [7]  
 ICM: C07H021-04  
 EXF ICS: C07K014-47  
 536/23.1; 536/23.5; 530/350  
 CAS INDEXING IS AVAILABLE FOR THIS PATENT.  
 L5 ANSWER 6 OF 16 USPTFLL  
 AN 2001:158014 USPTFLL  
 TI Methods, software and apparatus for identifying genomic regions harboring  
 a gene associated with a detectable trait  
 IN Schork, Nicholas J., Shaker Heights, OH, United States  
 Essloux, Laurent, Paris, France  
 Cohen-Akenine, Annick, Paris, France  
 Blumenfeld, Marta, Paris, France  
 Cohen, Daniel, Neuilly-sur-Seine, France  
 PA Genset, Paris, France (non-U.S. corporation)  
 PI US 6291182 B1 20010918  
 AI US 1999-438016 19991110 (9)  
 PRAI US 1998-107986P 19981110 (60)  
 DT US 1999-140785P 19990623 (60)  
 FS Utility  
 GRANTED

LN CNT 4241  
 INCL INCLM: 435/006.000  
 INCLM: 702/027.000  
 NCLM: 435/006.000  
 NCLM: 702/027.000  
 IC [7]  
 ICM: C12Q001-68  
 EXF ICS: G01N031-00  
 435/6; 530/300; 530/324; 530/325; 530/326; 712/200; 702/27  
 CAS INDEXING IS AVAILABLE FOR THIS PATENT.  
 L5 ANSWER 7 OF 16 USPTFLL  
 AN 2001:117151 USPTFLL  
 TI Prostate cancer gene  
 IN Cohen, Daniel, Neuilly sur Seine, France  
 Blumenfeld, Marta, Paris, France  
 Chumakov, Ilya, Vaux-le-Penil, FRANCE  
 Bougueleret, Lydie, Vanves, France  
 PA Genset, France (non-U.S. corporation)  
 PI US 6265546 B1 20010724  
 AI US 1999-338907 19990623 (9)  
 RLI Continuation-in-part of Ser. No. US 1998-218207, filed on 22 Dec 1998  
 now patented, Pat. No. US 5945522  
 PRAI US 1998-99658P 19980909 (60)  
 DT US 1998-107986P 19981110 (60)  
 FS Utility  
 GRANTED  
 LN CNT 7782  
 INCL INCLM: 530/350.000  
 INCLM: 424/174.100; 435/007.100; 514/002.000  
 NCLM: 530/350.000  
 NCLM: 424/174.100; 435/007.100  
 IC [7]  
 ICM: C07K001-00  
 EXF ICS: A61K039-395; A01N037-18; G01N035-53  
 530/350; 530/300; 530/324; 530/325; 514/2; 514/10; 514/12; 424/174.1;  
 435/7.1  
 CAS INDEXING IS AVAILABLE FOR THIS PATENT.  
 L5 ANSWER 8 OF 16 USPTFLL  
 AN 2001:86218 USPTFLL  
 TI Sequence specific DNA binding p53  
 IN Vogelstein, Bert, Baltimore, MD, United States  
 Kinzler, Kenneth W., Belair, MD, United States  
 Sherman, Michael I., Allendale, NJ, United States  
 PA The Johns Hopkins University, Baltimore, MD, United States (U.S.  
 corporation)  
 PI US 6245515 B1 20010612  
 AI US 1999-399773 19990921 (9)  
 RLI Division of Ser. No. US 1994-299074, filed on 1 Sep 1994, now patented,  
 Pat. No. US 5955263 Division of Ser. No. US 1992-860758, filed on 31 Mar  
 1992, now patented, Pat. No. US 5362623 Continuation-in-part of Ser. No.  
 US 1991-715182, filed on 14 Jun 1991, now abandoned  
 DT Utility  
 FS GRANTED  
 LN CNT 1804  
 INCL INCLM: 435/006.000  
 INCLM: 536/023.100; 536/024.330; 536/024.500; 536/025.300; 536/027.000  
 NCLM: 435/006.000  
 NCLM: 536/023.100; 536/024.330; 536/024.500; 536/025.300  
 IC [7]  
 ICM: C12Q001-68

**THIS PAGE BLANK (USPTO)**

AU 9915/40	A1	19990/12	AU 1999-15/40	19981222
EP 991770	A2	20000412	EP 1998-960061	19981222
EP 991770	B1	20001129		

EP 991770	B1	20001129	EP 1998-960061	19981222
EP 991770	A1	19980712	AU 1999-15740	19981222
AU 9915740	AA	19980701	CA 1998-231146	19981222
US 5945522	A	19990831	US 1997-9996306	19971222
CA 231146	AA	19980701	CA 1998-231146	19981222
EP 991770	B1	20001129	EP 1998-960061	19981222

**THIS PAGE BLANK (USPTO)**

R: AT, BE, CH, DE, DK, ES, FR, GB, GR, IT, LI, LU, NL, SE, MC, PT,  
 IE, FI  
 EP 1052292 A1 20001115 EP 2000-115991 19981222  
 R: AT, BE, CH, DE, DK, ES, FR, GB, GR, IT, LI, LU, NL, SE, MC, PT,  
 IE, FI  
 AT 197817 E 20001215 AT 1998-960061 19981222  
 US 2002516657 T2 20020611 JP 2000-525562 19981222  
 PRAI US 1997-996306 A 19971222  
 US 1998-99658P P 19980909  
 EP 1998-960061 A3 19981222  
 WO 1998-1B2133 W 19981222

ANSWER 13 OF 16 USPTAFULL  
 LN 1999-113551 USPTAFULL  
 TI Sequence specific DNA binding by p53  
 IN Vogelstein, Bert, Baltimore, MD, United States  
 Kinzler, Kenneth W., Baltimore, MD, United States  
 Sherman, Michael I., Glen Ridge, NJ, United States  
 Johns Hopkins University, Baltimore, MD, United States (U.S. corporation)  
 PA Genzyme Corporation, Framingham, MA, United States (U.S. corporation)  
 PI US 5955263 19990921  
 AI US 1994-299074 19940901 (8)  
 RLI Division of Ser. No. US 1992-860758, filed on 31 Mar 1992, now patented, Pat. No. US 5362623 which is a continuation-in-part of Ser. No. US 1991-715182, filed on 14 Jun 1991, now abandoned  
 DT Utility  
 FS Granted  
 LN.CNT 2012  
 INCL INCLM: 435/006.000  
 INCLS: 435/007.230; 436/063.000; 436/064.000; 536/023.100; 536/024.100  
 NCL NCLM: 435/006.000  
 IC [6]  
 ICM: C120001-68  
 ICS: G01N033-574; G01N033-48; C07H021-02  
 EXF 435/6; 435/7.23; 435/7.92; 536/23.5; 536/23.1; 536/24.1; 436/63; 436/64  
 CAS INDEXING IS AVAILABLE FOR THIS PATENT.

ANSWER 14 OF 16 USPTAFULL  
 LN 1999-102911 USPTAFULL  
 TI Prostate cancer gene  
 IN Cohen, Daniel, Fontenay-sous-Bois, France  
 Chumakov, Ilya, Vaux-le-Penil, France  
 Blumenfeld, Marta, Paris, France  
 Bouguetel, Lydie, Vanves, France  
 PA GENSET, Paris, France (non-U.S. corporation)  
 PI US 5945522 19990831  
 AI US 1997-996306 19971222 (8)  
 DT Utility  
 FS Granted  
 LN.CNT 5177  
 INCL INCLM: 536/023.100  
 INCLS: 435/006.000; 536/024.100; 536/024.310; 536/024.320; 536/024.330;  
 NCL NCLM: 536/023.100  
 IC NCLS: 435/006.000; 536/024.100; 536/024.300; 536/024.310; 536/024.320;  
 536/024.330  
 IC [6]  
 ICM: C07H021-02  
 ICS: C07H021-04; C120001-68  
 EXF 435/6; 536/24.3; 536/24.31; 536/24.32; 536/24.33; 536/24.1; 536/23.1  
 CAS INDEXING IS AVAILABLE FOR THIS PATENT.

ANSWER 15 OF 16 USPTAFULL  
 LN 97-51869 USPTAFULL  
 TI Isolated nucleic acid encoding a ubiquitous nuclear receptor  
 IN Liao, Shutsung, Chicago, IL, United States  
 Song, Ching, Durham, NC, United States  
 Arch Development Corporation, Chicago, IL, United States (U.S. corporation)  
 PA US 5639616 19970617  
 AI US 1994-342411 19941118 (8)  
 RLI Continuation-in-part of Ser. No. US 1993-152003, filed on 10 Nov 1993, now abandoned  
 DT Utility  
 FS Granted  
 LN.CNT 4472  
 INCL INCLM: 435/007.100  
 INCLS: 435/069.100; 435/252.300; 435/320.100; 536/023.500; 536/024.300  
 NCL NCLM: 435/007.100  
 IC NCLS: 435/069.100; 435/252.300; 435/320.100; 536/023.500; 536/024.300  
 IC [6]  
 ICM: C12N005-10  
 ICS: C12N015-12  
 EXF 435/67.1; 435/69.1; 435/252.3; 435/320.1; 536/23.5; 536/24.3  
 CAS INDEXING IS AVAILABLE FOR THIS PATENT.

ANSWER 16 OF 16 USPTAFULL  
 LN 94-97462 USPTAFULL  
 TI Sequence specific DNA binding by p53  
 IN Vogelstein, Bert, Baltimore, MD, United States  
 Kinzler, Kenneth W., Baltimore, MD, United States  
 Sherman, Michael I., Glen Ridge, NJ, United States  
 The John Hopkins University, Baltimore, MD, United States (U.S. corporation)  
 PA Pharmacia, Inc., Allendale, NJ, United States (U.S. corporation)  
 PI US 5362623 19941108  
 AI US 1992-860758 19920331 (7)  
 RLI Continuation-in-part of Ser. No. US 1991-715182, filed on 14 Jun 1991  
 DT Utility  
 FS Granted  
 LN.CNT 1741  
 INCL INCLM: 435/006.000  
 INCLS: 435/320.100; 536/024.100; 536/024.310; 424/002.000  
 NCL NCLM: 435/006.000  
 IC NCLS: 435/320.100; 536/024.100; 536/024.310  
 IC [5]  
 ICM: C120001-68  
 ICS: 435/5; 435/6; 435/7.1; 435/7.23; 435/7.92; 435/7.94; 435/70.3;  
 435/272; 435/975; 435/320.1; 436/501; 436/518; 436/63; 436/64; 436/811;  
 436/813; 424/2; 536/23.1; 536/24.31; 536/24.1  
 EXF 435/5; 435/6; 435/7.1; 435/7.23; 435/7.92; 435/7.94; 435/70.3;  
 435/272; 435/975; 435/320.1; 436/501; 436/518; 436/63; 436/64; 436/811;  
 436/813; 424/2; 536/23.1; 536/24.31; 536/24.1  
 CAS INDEXING IS AVAILABLE FOR THIS PATENT.

**THIS PAGE BLANK (USPTO)**



GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 28, 2002, 11:28:33 ; Search time 67.07 Seconds  
(without alignments)  
938.872 Million cell updates/sec

Title: US-09-853-526-5

Sequence: 1 MLSTLVHTYMRRLPSVY.....YVNWVITGTLGLAWTIKA 364

Scoring table: OLIGO  
Gapop 60.0 , Gapext 60.0

Searched: 562222 seqs, 172994929 residues

Word size : 0

Total number of hits satisfying chosen parameters: 562222

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : SPTREMBL.19:\*

1: sp\_archaea:\*

2: sp\_bacteria:\*

3: sp\_fungi:\*

4: sp\_human:\*

5: sp\_invertebrate:\*

6: sp\_mammal:\*

7: sp\_mhc:\*

8: sp\_organelle:\*

9: sp\_phage:\*

10: sp\_plant:\*

11: sp\_rodent:\*

12: sp\_virus:\*

13: sp\_vertebrate:\*

14: sp\_unclassified:\*

15: sp\_virus:\*

16: sp\_bacteriaph:\*

17: sp\_archaeap:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	364	100.0	364	4	Q9BQG4	Q9BQG4 homo sapien
2	252	69.2	353	4	Q9NUO2	Q9NUO2 homo sapien
3	32	8.8	354	11	Q9DIB8	Q9DIB8 mus musculu
4	9	2.5	264	16	Q9PQET	Q9PQET ureaplasma
5	9	2.5	185	5	Q9SR12	Q9SR12 caenorhabdi
6	8	2.2	115	16	Q9SE87	Q9SE87 rhizobium 1
7	8	2.2	236	16	Q9A675	Q9A675 caulobacter
8	8	2.2	262	16	Q9A743	Q9A743 chlamydia t
9	8	2.2	282	16	Q9A380	Q9A380 treponema p
10	8	2.2	411	16	Q9A4V3	Q9A4V3 caulobacter
11	8	2.2	430	2	Q9A517	Q9A517 acinetobact
12	8	2.2	454	3	Q9PSW2	Q9PSW2 neurospora
13	8	2.2	458	16	Q9KFI0	Q9KFI0 bacillus ha
14	8	2.2	459	16	Q9A503	Q9A503 bacillus su
15	8	2.2	483	10	Q9A927	Q9A927 pismu saliv
16	8	2.2	555	5	Q961R9	Q961R9 drosophila

17	8	2.2	557	5	Q9VEX8	Q9VEX8 drosophila
18	8	2.2	1820	5	Q9VVG0	Q9VVG0 drosophila
19	7	1.9	54	6	Q28219	Q28219 cercopithec
20	7	1.9	70	5	Q9NE57	Q9NE57 drosophila
21	7	1.9	78	16	Q9BRC0	Q9BRC0 rhizobium 1
22	7	1.9	116	8	Q9ATR2	Q9ATR2 neoscoelus
23	7	1.9	126	11	Q62370	Q62370 mus musculu
24	7	1.9	132	11	Q9D581	Q9D581 mus musculu
25	7	1.9	137	12	Q55581	Q55581 leucania se
26	7	1.9	141	6	Q28218	Q28218 cercopithec
27	7	1.9	143	16	Q9RRB9	Q9RRB9 mycoplasma
28	7	1.9	148	2	Q9RL25	Q9RL25 streptomyce
29	7	1.9	148	16	Q9CH79	Q9CH79 lactococcus
30	7	1.9	150	2	Q9AGV0	Q9AGV0 brucella ab
31	7	1.9	162	10	Q9LEP0	Q9LEP0 arabidopsis
32	7	1.9	163	2	Q9EX02	Q9EX02 streptomyce
33	7	1.9	170	5	Q9W4S6	Q9W4S6 drosophila
34	7	1.9	175	2	Q9EX12	Q9EX12 klebsiella
35	7	1.9	184	11	Q9CWC3	Q9CWC3 mus musculu
36	7	1.9	195	16	Q9KCD7	Q9KCD7 bacillus ha
37	7	1.9	196	10	Q9LH8	Q9LH8 pseudomonas
38	7	1.9	209	16	Q9I0U7	Q9I0U7 pseudomonas
39	7	1.9	209	16	Q9CJKA	Q9CJKA pasteurella
40	7	1.9	211	16	Q67841	Q67841 aquifex aeo
41	7	1.9	220	2	Q9EWV7	Q9EWV7 streptomyce
42	7	1.9	225	16	Q97NE9	Q97NE9 streptococc
43	7	1.9	230	16	Q9PHZ5	Q9PHZ5 campylobact
44	7	1.9	234	2	Q32330	Q32330 clostridium
45	7	1.9	240	12	Q9WHT2	Q9WHT2 thailand to

## ALIGNMENTS

RESULT 1

Q9BQG4 PRELIMINARY; PRT: 364 AA.

Q9BQG4: 01-JUN-2001 (TREMBLrel. 17, Created)

DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)

DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)

DE HYPOTHETICAL 42.1 KDA PROTEIN (LYSOPHOSPHATIDIC ACID

DE ACYLTRANSFERASE-EPSILON) (EC 2.3.1.51).

GN DKFPZP761C222 OR LPAAT-E.

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

OX NCBI\_TaxID=9606;

RP [1]

RP SEQUENCE FROM N.A.

RC TISSUE=AMYGDALA;

RX MEDLINE=21154917; PubMed=11230166;

RA Wiemann S., Well B., Wellenreuther R., Gassenhuber J., Glasel S.,

RA Anstorge W., Boecker M., Bloeker H., Bauersachs S., Blum H.,

RA Lauber J., Duesterhoeft B., Beyer A., Koehler K., Strack N.,

RA Mewes H.W., Oltenswaelder B., Obermaier B., Tampe J., Heubner D.,

RA Wambutt R., Korn B., Klein M., Poustka A.;

RA "Towards a Catalog of Human Genes and Proteins: Sequencing and

RT Analysis of 500 Novel Complete Protein Coding Human cDNAs."

RL Genome Res. 11:422-435(2001).

RN [2]

RP SEQUENCE FROM N.A.

RA Leung D.W.;

RT "Cloning and expression of LPAAT-epsilon."

RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.

DR EMBL: AL136587; CAB65522.1; "

DR EMBL: AF375789; AAK54809.1; "

DR InterPro: IPR002123; Acyltransferase.

DR Pfam: PF01535; Acyltransferase; 1.

DR Hypothetical protein; Transferase; Acyltransferase.

SO SEQUENCE 364 AA; 42072 MW; 90A0F87FC7C78081 CRC64;

Query Match 100.0%; Score 364; DB 4; Length 364;  
 Best Local Similarity 100.0%; Pred. No. 0;  
 Matches 364; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MLSTVLTHTYMRVLLPSVLLGTAFTYVLAAGVWRLSFLPARFYQALDDRLCYGOS 60  
 DB 1 MLSTVLTHTYMRVLLPSVLLGTAFTYVLAAGVWRLSFLPARFYQALDDRLCYGOS 60  
 QY 61 MVLFFENYTGVOITLYGDDLPKKNENIYLANHSTVDMIVADILAIROMALGHVYVVK 120  
 DB 61 MVLFFENYTGVOITLYGDDLPKKNENIYLANHSTVDMIVADILAIROMALGHVYVVK 120  
 QY 121 EGLWMLPLCYGCFYFAGHGIYVKSRAKFNKEMENKLOSVDYDAGTDMYLVIFPEGTNYNE 180  
 DB 121 EGLWMLPLCYGCFYFAGHGIYVKSRAKFNKEMENKLOSVDYDAGTDMYLVIFPEGTNYNE 180  
 QY 181 QTKVLSAQFAAORGIAVLAKHVLTPRIKATHVAFDCMKNYLDIAYDTVVEGKDDGQ 240  
 DB 181 QTKVLSAQFAAORGIAVLAKHVLTPRIKATHVAFDCMKNYLDIAYDTVVEGKDDGQ 240  
 QY 241 RRSPTWTEFLCKECPKIHIDRIKDDVPEOEHRMRWLHERPEIKDKMLIEFESPD 300  
 DB 241 RRSPTWTEFLCKECPKIHIDRIKDDVPEOEHRMRWLHERPEIKDKMLIEFESPD 300  
 QY 301 PERKRPFGSVNSKISTIKTLPMSLISGLTAGMLMTDAGRKLVTWITYGTLGCLMV 360  
 DB 301 PERKRPFGSVNSKISTIKTLPMSLISGLTAGMLMTDAGRKLVTWITYGTLGCLMV 360  
 QY 361 TTKA 364  
 DB 361 TTKA 364

RESULT 2  
 Q9NUO2 PRELIMINARY; PRT: 353 AA.  
 AC Q9NUO2;  
 DT 01-OCT-2000 (Tremblrel. 15, Created)  
 DT 01-OCT-2000 (Tremblrel. 15, Last sequence update)  
 DE 01-JUN-2001 (Tremblrel. 17, Last annotation update)  
 DE CDNA FLJ11210 F1S, CLONE PLACE1007954.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniala; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 NCBI\_Taxid=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=PLACENTA.  
 RA Iogel T., Ota T., Hayashi K., Sugiyama T., Otsuki T., Suzuki Y.,  
 RA Nishikawa T., Nagai K., Sugano S., Shiretori A., Sudo H.,  
 RA Wagaatsuma M., Hosoi T., Kaku Y., Kodaira H., Kondo H., Sugawara M.,  
 RA Wakatashi M., Chida Y., Ishida S., Murakawa K., Ono Y., Takiguchi S.,  
 RA Yamamoto J., Wakamatsu K., Nakamura Y., Ishii S., Kawai Y., Saito K.,  
 RA Niinomiya K., Iwayanagi T.,  
 RT "MDO human cDNA sequencing project."  
 RL Submitted (FEB-2000) to the EMBL/Genbank/DBJ databases.  
 DR EMBL: AK002072; BAA92069.1;  
 DR InterPro: IPR002123; Acyltransferase.  
 DR Pfam: PF01553; Acyltransferase; 1.  
 SO SEQUENCE 353 AA; 40799 MW; 0A8A14A2361F1B14 CRC64;

Query Match 69.2%; Score 252; DB 4; Length 353;  
 Best Local Similarity 99.7%; Pred. No. 1,4e-250;  
 Matches 352; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 12 MYRLPSVLLTAPTYVLAAGVWRLSFLPARFYQALDDRLCYGOSVWVLEFFENITG 71  
 DB 1 MYRLPSVLLTAPTYVLAAGVWRLSFLPARFYQALDDRLCYGOSVWVLEFFENITG 71  
 QY 72 VOILLYGDLPKKNENIYLANHSTVDMIVADILAIROMALGHVYVVKELKMLPLYGC 131  
 DB 72 VOILLYGDLPKKNENIYLANHSTVDMIVADILAIROMALGHVYVVKELKMLPLYGC 131

DB 61 VOILLYGDLPKKNENIYLANHSTVDMIVADILAIROMALGHVYVVKELKMLPLYGC 120  
 QY 133 YFAOHGIIYVKSRAKFNKEMENKLOSVDYDAGTDMYLVIFPEGTNYNEPOTKYLSSAQF 191  
 DB 121 YFAOHGIIYVKSRAKFNKEMENKLOSVDYDAGTDMYLVIFPEGTNYNEPOTKYLSSAQF 180  
 QY 192 AAORGLAVLKHVLTPIKATHVAFDCMKNYLDIAYDTVVEGKDDGQRRSPWTEFL 251  
 DB 181 AAORGLAVLKHVLTPIKATHVAFDCMKNYLDIAYDTVVEGKDDGQRRSPWTEFL 240  
 QY 252 CKCEPKIHIDRIKDDVPEOEHRMRWLHERPEIKDKMLIEFESSPPERKRPFGS 311  
 DB 241 CKCEPKIHIDRIKDDVPEOEHRMRWLHERPEIKDKMLIEFESSPPERKRPFGS 300  
 QY 312 VNSKLSIKTLPMSLISGLTAGMLMTDAGRKLVTWITYGTLGCLMTTKA 364  
 DB 301 VNSKLSIKTLPMSLISGLTAGMLMTDAGRKLVTWITYGTLGCLMTTKA 353

RESULT 3  
 Q9DIE8 PRELIMINARY; PRT: 354 AA.  
 AC Q9DIE8;  
 DT 01-JUN-2001 (Tremblrel. 17, Created)  
 DT 01-JUN-2001 (Tremblrel. 17, Last sequence update)  
 DE 01-JUN-2001 (Tremblrel. 17, Last annotation update)  
 DE 1110013A05RIK PROTEIN.  
 GN 1110013A05RIK.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniala; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 NCBI\_Taxid=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=C57BL/6J; TISSUE=EMBRYO;  
 RX MEDLINE=24085660; PubMed=11217851;  
 RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,  
 RA Arakawa T., Hara A., Fukunishi Y., Kono H., Adachi J., Fukuda S.,  
 RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamana I.,  
 RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,  
 RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,  
 RA Felschmann W., Gaasterland T., Gissi C., King B., Kochia H.,  
 RA Kuehl P., Lewis S., Matsuo Y., Nakado I., Pesole G., Quackenbush J.,  
 RA Schriml L.M., Straubli F., Suzuki R., Tomita M., Wagner L., Mashio T.,  
 RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,  
 RA Blake J., Boffelli D., Bojunga N., Carninci P., Fujita M., Gariboldi M.,  
 RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,  
 RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,  
 RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Womdearts P.,  
 RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,  
 RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-P.,  
 RA Suzuki H., Toyooka K.A., Wang K.H., Weitz C., Whitaker C., Wilmink L.,  
 RA Wyszynski-Boris A., Yoshida K., Hasegawa Y., Kawai H., Kohlski S.,  
 RA Hayashizaki Y.,  
 RT "Functional annotation of a full-length mouse cDNA collection."  
 RL Nature 409:685-690(2001).  
 DR EMBL: AK003649; BAB22915.1;  
 DR MGD: MGI:1915880; 1110013A05RIK.  
 DR InterPro: IPR002123; Acyltransferase.  
 DR Pfam: PF01553; Acyltransferase; 1.  
 SO SEQUENCE 354 AA; 40943 MW; 447EEF924B91E800 CRC64;

Query Match 8.8%; Score 32; DB 11; Length 354;  
 Best Local Similarity 100.0%; Pred. No. 3.1e-24;  
 Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 185 LSASQAFAAORGIAVLAKHVLTPRIKATHVAFD 216  
 DB 174 LSASQAFAAORGIAVLAKHVLTPRIKATHVAFD 205

RESULT 4

```

Q9POE7
ID Q9POE7 PRELIMINARY; PRT; 264 AA.
AC Q9POE7;
DT 01-OCT-2000 (TREMBLrel. 15, Created)
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
DE 1-ACYL-SN-GLYCEROL-3-PHOSPHATE ACYLTRANSFERASE.
GN PLSC OR U0344.
OS Ureaplasma parvum (Ureaplasma urealyticum biotype 1).
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Mollicutes;
OC Mycoplasmataceae; Ureaplasma.
OX NCBI_TaxID=134821;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-SEROVAR 3:
RX MEDLINE=20500219; PubMed=11048724;
RA Glass J.I., Lefkowitz E.J., Glass J.S., Helner C.R., Chen E.Y.,
RA Cassell G.H.;
RT "The complete sequence of the mucosal pathogen Ureaplasma
RT urealyticum."
RL Nature 407:757-762(2000).
DR EMBL: AE002131; AAF30753.1; -
DR InterPro: IPR002123; Acyltransferase.
DR Pfam: PF01553; Acyltransferase; 1.
KW Complete proteome.
SQ SEQUENCE 264 AA; 30687 MW; DDC5727A07C74F7 CRC64;

Query Match 2.5%; Score 9; DB 16; Length 264;
Best Local Similarity 100.0%; Pred. No. 1.1;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 168 LVIFPEGTR 176
Db 156 LVIFPEGTR 164

RESULT 5
Q95RI2 PRELIMINARY; PRT; 386 AA.
ID Q95RI2;
AC Q95RI2;
DT 01-DEC-2001 (TREMBLrel. 19, Created)
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE HYPOTHETICAL 44.0 KDA PROTEIN.
GN F28B3.9.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-BRISTOL N2;
RX MEDLINE=99069613; PubMed=9851916;
RA None;
RT "Genome sequence of the nematode C. elegans: a platform for
RT investigating biology. The C. elegans Sequencing Consortium."
RL Science 282:2012-2018(1998).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-BRISTOL N2;
RA Geisel C., Kramer J., Smith A.;
RT "The sequence of C. elegans cosmid F28B3."
RL Submitted (May-1997) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN-BRISTOL N2;
RA Waterston R.;
RT "Direct Submission."
RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF003136; AAK93853.1; -
KW Hypothetical protein.
SQ SEQUENCE 386 AA; 43978 MW; A1E3020C201452A6 CRC64;

```

```

Query Match 2.5%; Score 9; DB 5; Length 386;
Best Local Similarity 100.0%; Pred. No. 1.5;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 222 LDATVDVT 230
Db 230 LDATVDVT 238

RESULT 6
Q98E87 PRELIMINARY; PRT; 115 AA.
ID Q98E87;
AC Q98E87;
DT 01-OCT-2001 (TREMBLrel. 18, Created)
DT 01-OCT-2001 (TREMBLrel. 18, Last sequence update)
DT 01-OCT-2001 (TREMBLrel. 18, Last annotation update)
DE MLI4358 PROTEIN.
GN MLI4358.
OS Rhizobium loti (Mesorhizobium loti).
OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
OC Phyllobacteriaceae; Mesorhizobium.
OX NCBI_TaxID=381;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-MAFR303099;
RX MEDLINE=21082930; PubMed=11214968;
RA Kaneko T., Nakamura Y., Sato S., Asamizu E., Kato T., Sasamoto S.,
RA Watanabe A., Iidesawa K., Ishikawa K., Kawashima K., Kimura T.,
RA Kishida Y., Kiyokawa C., Kohara M., Matsumoto M., Matsuno A.,
RA Mochizuki Y., Nakayama S., Nakazaki N., Shimo S., Sugimoto M.,
RA Takeuchi C., Yamada M., Tabata S.;
RT "Complete genome structure of the nitrogen-fixing symbiotic bacterium
RT Mesorhizobium loti."
RL DNA Res. 7:331-338(2000).
DR EMBL: AP003004; BAB51033.1; -
KW Complete proteome.
SQ SEQUENCE 115 AA; 12335 MW; 6AF43D6451F42BE8 CRC64;

Query Match 2.2%; Score 8; DB 16; Length 115;
Best Local Similarity 100.0%; Pred. No. 5.6;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 185 LSASQAF 192
Db 31 LSASQAF 38

RESULT 7
Q9A675 PRELIMINARY; PRT; 236 AA.
ID Q9A675;
AC Q9A675;
DT 01-JUN-2001 (TREMBLrel. 17, Created)
DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE 1-ACYL-SN-GLYCEROL-3-PHOSPHATE ACYLTRANSFERASE.
GN CC2219.
OS Caulobacter crescentus.
OC Bacteria; Proteobacteria; alpha subdivision; Caulobacter group;
OC Caulobacter.
OX NCBI_TaxID=69394;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-ATCC 19089 / CB15;
RX MEDLINE=21173698; PubMed=11259647;
RA Nierman W.C., Feldblyum T.V., Laub M.T., Paulsen I.T., Nelson K.E.,
RA Eisen J., Heidelberg J.F., Alley M.R.K., Ohta N., Maddock J.R.,
RA Potocka I., Nelson W.C., Newton A., Stephens C., Phadke N.D., Ely B.,
RA Deboy R.T., Dodson R.J., Durkin A.S., Gwinn M.L., Haft D.H.,
RA Kolonay J.F., Smit J., Craven M.B., Khouri H., Shetty J., Berry K.,
RA Utterback T., Tran K., Wolf A., Vamthayan J., Ermolaeva M., White O.,

```

RA Salberg S.L., Venter J.C., Shapiro L., Fraser C.M.;  
 RT "Complete genome sequence of *Caulobacter crescentus*."  
 RL Proc. Natl. Acad. Sci. U.S.A. 98:4136-4141(2001).  
 DR EMBL: AE005893; AAK24190.1; -  
 DR TIGR: CC2219; -  
 DR InterPro: IPR002123; Acyltransferase.  
 DR Pfam: PF01553; Acyltransferase; 1.  
 KW Transferase; Acyltransferase; Complete proteome.  
 SQ SEQUENCE 236 AA; 26382 MW; 85694D396B8CA1A0 CRC64;

Query Match 2.2%; Score 8; DB 16; Length 236;  
 Best Local Similarity 100.0%; Pred. No. 11;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 169 VIPEPCTR 176  
 DB 146 VIPEPCTR 153

RESULT 8  
 084743 PRELIMINARY; PRT; 262 AA.  
 AC 084743:  
 DT 01-NOV-1998 (TREMBlrel. 08, Created)  
 DT 01-NOV-1998 (TREMBlrel. 08, Last sequence update)  
 DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)  
 DE METAL DEPENDENT HYDROLASE.  
 GN YJCJ OR CT738.  
 OS Chlamydia trachomatis.  
 OC Bacteria; Chlamydiales; Chlamydiaceae; Chlamydia.  
 OX NCBI\_Taxid=813;  
 RN (1)  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=D/UW-3/CX;  
 RX MEDLINE=99000809; PubMed=9784136;  
 RA Stephens R.S., Kaiman S., Lammel C.J., Fan J., Marathe R., Aravind L.,  
 RA Mitchell W.P., Olinger L., Tatusov R.L., Zhao Q., Koonin E.V.,  
 RA Davis R.W.;  
 RT "Genome sequence of an obligate intracellular pathogen of humans:  
 RT Chlamydia trachomatis."  
 RL Science 282:754-759(1998).  
 DR EMBL: AE001344; AAC68333.1; -  
 KW Hydrolase; Complete proteome.  
 SQ SEQUENCE 262 AA; 29082 MW; E36A6298A23CEFBF CRC64;

Query Match 2.2%; Score 8; DB 16; Length 262;  
 Best Local Similarity 100.0%; Pred. No. 12;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 202 HVLTPRIK 209  
 DB 202 HVLTPRIK 209

RESULT 9  
 083380 PRELIMINARY; PRT; 282 AA.  
 AC 083380:  
 DT 01-NOV-1998 (TREMBlrel. 08, Created)  
 DT 01-NOV-1998 (TREMBlrel. 08, Last sequence update)  
 DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)  
 DE LYSOPHOSPHATIDIC ACID ACYLTRANSFERASE, PUTATIVE.  
 GN TP0361.  
 OS Treponema pallidum.  
 OC Bacteria; Spirochaetales; Spirochaetaceae; Treponema.  
 OX NCBI\_Taxid=160;  
 RN (1)  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=NICHOLS;  
 RX MEDLINE=98332770; PubMed=9665876;  
 RA Fraser C.M., Norris S.J., Weinstock G.M., White O., Sutton G.G.,

RA Dodson R., Gwinn M., Hickey E.K., Clayton R., Ketchum K.A.,  
 RA Sodergren E., Harsham J.M., McLeod J.P., Salberg S., Peterson J.,  
 RA Khaliq H., Richardson D., Howell J.K., Chidambaram M., Uterback T.,  
 RA McDonald L., Artlich P., Bowman C., Cotton M.D., Fujii C., Garland S.,  
 RA Hatch B., Horst K., Roberts K., Sandusky M., Weidman J., Smith H.O.,  
 RA Venter J.C.;  
 RT "Complete genome sequence of *Treponema pallidum*, the syphilis  
 RT spirochete."  
 RL Science 281:375-388(1998).  
 DR EMBL: AE001215; AAC65346.1; -  
 DR TIGR: TP0361; -  
 DR InterPro: IPR002123; Acyltransferase.  
 DR Pfam: PF01553; Acyltransferase; 1.  
 KW Transferase; Acyltransferase; Complete proteome.  
 SQ SEQUENCE 282 AA; 31700 MW; DEF69003CD6C1AD CRC64;

Query Match 2.2%; Score 8; DB 16; Length 282;  
 Best Local Similarity 100.0%; Pred. No. 12;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 169 VIPEPCTR 176  
 DB 170 VIPEPCTR 177

RESULT 10  
 09A4V3 PRELIMINARY; PRT; 411 AA.  
 AC 09A4V3:  
 DT 01-JUN-2001 (TREMBlrel. 17, Created)  
 DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)  
 DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)  
 DE METAL ION EFFLUX MEMBRANE FUSION PROTEIN FAMILY.  
 GN CC2722.  
 OS *Caulobacter crescentus*.  
 OC Bacteria; Proteobacteria; alpha subdivision; Caulobacter group;  
 OC Caulobacter.  
 OX NCBI\_Taxid=69394;  
 RN (1)  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=ATCC 19089 / CH15;  
 RX MEDLINE=21173698; PubMed=11259647;  
 RA Nierman W.C., Feldblyum T.V., Laub M.T., Paulsen I.T., Nelson K.E.,  
 RA Eisen J., Heidelberg J.F., Alley M.R.K., Ohta N., Maddock J.R.,  
 RA Potocka I., Nelson W.C., Newton A., Stephens C., Phadke N.D., Ely B.,  
 RA DeBoy R.T., Dodson R.J., Durkin A.S., Gwinn M.L., Haft D.H.,  
 RA Kolonay J.F., Smit J., Craven M.B., Khouri H., Shetty J., Berry K.,  
 RA Uterback T., Tran K., Wolf A., Yamathavan J., Ermolaeva M., White O.,  
 RA Salberg S.L., Venter J.C., Shapiro L., Fraser C.M.;  
 RT "Complete genome sequence of *Caulobacter crescentus*."  
 RL Proc. Natl. Acad. Sci. U.S.A. 98:4136-4141(2001).  
 DR EMBL: AE005938; AAK24687.1; -  
 DR TIGR: CC2722; -  
 DR InterPro: IPR002215; H1YD.  
 DR Pfam: PF00529; H1YD; 1.  
 KW Complete proteome.  
 SQ SEQUENCE 411 AA; 41818 MW; 1BA7849B9BCFC22E CRC64;

Query Match 2.2%; Score 8; DB 16; Length 411;  
 Best Local Similarity 100.0%; Pred. No. 17;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 327 ILSGLTAG 334  
 DB 380 ILSGLTAG 387

RESULT 11  
 054517 PRELIMINARY; PRT; 430 AA.  
 AC 054517:  
 RX MEDLINE=98332770; PubMed=9665876;  
 RA Fraser C.M., Norris S.J., Weinstock G.M., White O., Sutton G.G.,

DT 01-JUN-1998 (TREMBLrel. 06, Created)  
DT 01-JUN-1998 (TREMBLrel. 06, Last sequence update)  
DT 01-OCT-2001 (TREMBLrel. 18, Last annotation update)  
DE 3-DEOXY-D-MANNO-2-OCTULOSONATE TRANSFERASE.  
GN MAA.  
OS Acinetobacter baumannii.  
OC Bacteria; Proteobacteria; gamma subdivision; Moraxellaceae;  
OC Acinetobacter.  
OX NCBI\_TaxID=470;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-ATCC 19606, AND ATCC 15308;  
RA Bode C.E., Brabetz W., Brade H.;  
RL Submitted (JAN-1998) to the EMBL/GenBank/DBJ databases.  
DR EMBL: 296926; CAB09652.1; -;  
KW EMBL: 296925; CAB09651.1; -;  
SQ TRANSFERASE.  
KW TRANSFERASE.  
SQ SEQUENCE 430 AA; 49070 MW; 695293748185D272 CRC64;

Query Match 2.2%; Score 8; DB 2; Length 430;  
Best Local Similarity 100.0%; Pred. No. 18;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 329 SGLTAGML 336  
Db 171 SGLTAGML 178

RESULT 12  
O9P5W2 PRELIMINARY; PRT; 454 AA.  
AC O9P5W2;  
DT 01-OCT-2000 (TREMBLrel. 15, Created)  
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)  
DE 01-DEC-2001 (TREMBLrel. 19, Last annotation update)  
DE RELATED TO GUANINE DEAMINASE.  
GN B3E4.200.  
OS Neurospora crassa.  
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;  
OC Sordariales; Sordariaceae; Neurospora.  
OX NCBI\_TaxID=5141;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Schulte U., Aign V., Hoheisel J., Brandt P., Fartmann B., Holland R.,  
RA Nyakatura G., Mewes H.W., Mannhaupt G.;  
RL Submitted (MAY-2000) to the EMBL/GenBank/DBJ databases.  
RN [2]  
RP SEQUENCE FROM N.A.  
RA German Neurospora genome project;  
RL Submitted (NOV-2001) to the EMBL/GenBank/DBJ databases.  
DR EMBL: AL355931; CAB91408.2; -;  
SQ SEQUENCE 454 AA; 49778 MW; 8C1EC14BAEC421D CRC64;

Query Match 2.2%; Score 8; DB 3; Length 454;  
Best Local Similarity 100.0%; Pred. No. 19;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 200 LKHVLPFR 207  
Db 208 LKHVLPFR 215

RESULT 13  
O9KF10 PRELIMINARY; PRT; 458 AA.  
AC O9KF10;  
DT 01-OCT-2000 (TREMBLrel. 15, Created)  
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)  
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)  
DE RNA METHYLTRANSFERASE.  
GN BH0687.

OS Bacillus halodurans.  
OC Bacteria; Firmicutes; Bacillus/Clostridium group;  
OC Bacillus/Staphylococcus group; Bacillus.  
OX NCBI\_TaxID=86665;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-C-125 / JCM 9153;  
RX MEDLINE-20512582; PubMed-11058132;  
RA Takami H., Nakasone K., Takaki Y., Maeno G., Sasaki R., Masui N.,  
RA Fujii F., Hirama C., Nakamura Y., Ogasawara N., Kuhara S.,  
RA Horikoshi K.;  
RT "Complete genome sequence of the alkaliphilic bacterium Bacillus  
RT halodurans and genomic sequence comparison with Bacillus subtilis.";  
RL Nucleic Acids Res. 28:4317-4331(2000).  
DR EMBL: AP001509; BAB04406.1; -;  
DR InterPro: IPR000051; SAM\_bind.  
DR InterPro: IPR001566; TRMA\_1.  
DR PROSITE: PS01230; TRMA\_1; UNKNOWN\_1.  
KW Transferase; Methyltransferase; Complete proteome.  
SQ SEQUENCE 458 AA; 51555 MW; 3D33A661E8891240 CRC64;

Query Match 2.2%; Score 8; DB 16; Length 458;  
Best Local Similarity 100.0%; Pred. No. 19;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 178 NPEOTKVL 185  
Db 291 NPEOTKVL 298

RESULT 14  
O31503 PRELIMINARY; PRT; 459 AA.  
ID O31503;  
DT 01-JAN-1998 (TREMBLrel. 05, Created)  
DT 01-JAN-1998 (TREMBLrel. 05, Last sequence update)  
DT 01-OCT-2001 (TREMBLrel. 18, Last annotation update)  
DE YEFA PROTEIN.  
GN YEFA.  
OS Bacillus subtilis.  
OC Bacteria; Firmicutes; Bacillus/Clostridium group;  
OC Bacillus/Staphylococcus group; Bacillus.  
OX NCBI\_TaxID=1423;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-168;  
RX MEDLINE-98044033; PubMed-9384377;  
RA Kunst F., Ogasawara N., Moser I., Albertini A.M., Alloni G.,  
RA Azevedo V., Bortner M.G., Bessieres P., Bolotin A., Borchert S.,  
RA Bortiss R., Boursier L., Brans A., Brunn M., Brignell S.C., Bron S.,  
RA Brouillet S., Brusch C.V., Caldwell B., Capiano V., Carter N.M.,  
RA Choi S.R., Codani J.J., Connerton I.F., Cummings N.J., Daniel R.A.,  
RA Denizot F., Devine K.M., Dusterhoft A., Ehrlich S.D., Emerson P.T.,  
RA Britian K.D., Errington J., Fabret C., Ferrari E., Fougere D.,  
RA Fritz C., Fujita M., Fujita Y., Funas S., Galizzi A., Galleron N.,  
RA Chim S.T., Glaser P., Goffeau A., Gollightly E.J., Grandi G.,  
RA Giusseppl G., Guy B.J., Haga K., Halech J., Harwood C.R., Henaut A.,  
RA Hilbert H., Holsappel S., Hosono S., Hullo M.F., Itaya M., Jones L.,  
RA Joris B., Karamata D., Kasahara Y., Klaerr-Blanchard M., Klein C.,  
RA Kobayashi Y., Koetter P., Koningsstein G., Krogh S., Kumano M.,  
RA Kurita K., Lapidus A., Lardinois S., Lauber J., Lazarevic V.,  
RA Lee S.M., Levine A., Liu H., Masuda S., Manuel C., Medigue C.,  
RA Medina N., Mellado R.P., Mizuno M., Moestl D., Nakai S., Noback M.,  
RA Noone D., O'Reilly M., Ogawa K., Ogihara A., Oudega B., Park S.H.,  
RA Paro V., Pohl T.M., Portetelle D., Portollik S., Prescott A.M.,  
RA Presecan E., Pujic P., Purnelle B., Rapoport G., Rey M., Reynolds S.,  
RA Rieger M., Rivolta C., Rocha E., Roche B., Rose M., Sadate Y.,  
RA Sato T., Scanlan E., Schleich S., Schroeter R., Scoffone F.,  
RA Sekiyuchi J., Sekowska A., Seror S.J., Seror P., Shin B.S., Soldo B.,  
RA Sorokin A., Tacconi E., Takagi T., Takahashi H., Takemaru K.,  
RA Takeuchi M., Tamakoshi A., Tanaka T., Terpstra P., Tognoni A.,  
RA Tosato V., Uchiyama S., Vandenbol M., Vannier F., Vassarotti A.,

RA Viart A., Wambutt R., Wedler E., Wedler H., Weitzenecker T.,  
 RA Winters P., Wipat A., Yamamoto H., Yamane K., Yasumoto K., Yata K.,  
 RA Yoshida K., Yoshikawa H.F., Zumbstein E., Yoshikawa H., Danchin A.,  
 RT "The complete genome sequence of the gram-positive bacterium *Bacillus*  
 RT *subtilis*.";  
 RL Mature 390:249-256(1997).

RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=168;  
 RA Kunst F., Ogasawara N., Yoshikawa H., Danchin A.,  
 RL Submitted (NOV-1997) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: Z99107; CAB12493.1; -;  
 DR InterPro: IPR002792; DJF90.  
 DR InterPro: IPR000051; SAM\_bind.  
 DR Pfam: PF01938; TRAM; 1.  
 DR PROSITE: PS01230; TRMA\_1; UNKNOWN\_1.  
 DR PROSITE: PS01231; TRMA\_2; 1.  
 KW Complete proteome.  
 SQ SEQUENCE 459 AA; 51819 MW; 35E99547779421D5 CRC64;

Query Match 2.2%: Score 8; DB 16; Length 459;  
 Best Local Similarity 100.0%: Pred. No. 19;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 178 NPEQTKVL 185  
 |||||  
 DB 292 NPEQTKVL 299

RESULT 15  
 ID 049927 PRELIMINARY; PRT; 483 AA.  
 AC 049927;  
 DT 01-JUN-1998 (TREMblrel. 06, Created)  
 DT 01-JUN-1998 (TREMblrel. 06, Last sequence update)  
 DT 01-DEC-2001 (TREMblrel. 19, Last annotation update)  
 DE P54 PROTEIN.  
 GN P54.  
 OS Pisum sativum (Garden pea).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
 OC eurosids I; Fabales; Fabaceae; Papilionoideae; Viciaeae; Pisum.  
 OX NCBI\_TaxID=3888;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Castillo J., Marquez J.A., Franco L., Ballestar E., Rodrigo M.I.,  
 RL Submitted (FEB-1997) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: Y11207; CAA72090.1; -;  
 DR HSSP: P50477; ICAU.  
 DR InterPro: IPR001113; Seedstore\_7s.  
 DR Pfam: PF00546; Seedstore\_7s; 1.  
 DR Pfam: PF02808; Seedstore\_7s\_C; 1.  
 DR PF02808; Seedstore\_7s\_C; 1.  
 SQ SEQUENCE 483 AA; 54662 MW; 8127BDAAA0178F3D CRC64;

Query Match 2.2%: Score 8; DB 10; Length 483;  
 Best Local Similarity 100.0%: Pred. No. 20;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 162 AGTPMYLV 169  
 |||||  
 DB 184 AGTPMYLV 191

Search completed: August 28, 2002, 11:28:34  
 Job time: 547 sec

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 28, 2002, 11:07:33 ; Search time 75.55 Seconds  
(without alignments)  
535.154 Million cell updates/sec

Title: US-09-853-526-5

Perfect score: 1924

Sequence: 1 MLSLVLTHTYSMRYLLPSPV.....YVNMIVYGTLLGLMTVIKA 364

Scoring table:

BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 747574 seqs, 111073796 residues

Total number of hits satisfying chosen parameters: 747574

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database :

A.Geneseq.032802:\*

1: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1980.DAT:\*  
2: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1981.DAT:\*  
3: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1982.DAT:\*  
4: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1983.DAT:\*  
5: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1984.DAT:\*  
6: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1985.DAT:\*  
7: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1986.DAT:\*  
8: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1987.DAT:\*  
9: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1988.DAT:\*  
10: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1989.DAT:\*  
11: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1990.DAT:\*  
12: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1991.DAT:\*  
13: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1992.DAT:\*  
14: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1993.DAT:\*  
15: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1994.DAT:\*  
16: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1995.DAT:\*  
17: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1996.DAT:\*  
18: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1997.DAT:\*  
19: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1998.DAT:\*  
20: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1999.DAT:\*  
21: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA2000.DAT:\*  
22: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA2001.DAT:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	1924	100.0	364	20	AAV36729	Human PGI protein
2	1924	100.0	364	22	AAAM0128	Human acyltransferase
3	1924	100.0	364	22	AAU00665	Human lysophosphatidyl transferase
4	1913	99.4	372	22	AAU41377	Human polyphosphatidyl transferase
5	1871	97.2	353	22	AAAB8532	Human secreted protein
6	1871	97.2	353	22	AAAB8552	Human secreted protein
7	1868	97.1	353	22	AAAB9359	Human protein sequence
8	1867	97.0	352	20	AAV36728	Human PGI protein
9	1861	96.7	353	22	AAAM39591	Human polyphosphatidyl transferase
10	1644	85.4	315	20	AAV36751	PGI splice variant
11	1568.5	81.5	354	20	AAV36741	Mouse PGI protein

12	1555.5	80.8	300	20	AAV36752	PGI splice variant
13	1498	77.9	291	20	AAV36744	PGI splice variant
14	1427	74.2	259	22	AAAB09653	Human gene 15 enco
15	1325	68.9	261	20	AAV36745	PGI splice variant
16	1203	62.5	228	20	AAV36740	Truncated PGI prot
17	1182.5	61.5	238	20	AAV36743	PGI splice variant
18	976	50.7	185	20	AAV36753	PGI splice variant
19	820	42.6	182	20	AAV36750	PGI splice variant
20	449	23.3	96	20	AAV36749	PGI splice variant
21	401.5	20.9	378	21	AAV36748	Arabidopsis thalia
22	396.5	20.6	352	21	AAV36747	Arabidopsis thalia
23	393.5	20.5	341	21	AAV36746	Arabidopsis thalia
24	390.5	20.3	375	21	AAV36743	Arabidopsis thalia
25	390.5	20.3	375	21	AAV36742	Arabidopsis thalia
26	387.5	20.1	351	21	AAV36735	Arabidopsis thalia
27	387.5	20.1	375	21	AAV36734	Arabidopsis thalia
28	387	20.1	257	22	AAV36730	Novel human diagno
29	387	20.1	980	22	AAV36727	Novel human diagno
30	384.5	20.0	351	21	AAV36726	Arabidopsis thalia
31	381	19.8	343	21	AAV36736	Arabidopsis thalia
32	381	19.8	343	21	AAV36735	Arabidopsis thalia
33	381	19.8	343	21	AAV36734	Arabidopsis thalia
34	330	17.2	66	20	AAV36748	PGI splice variant
35	329	17.1	68	20	AAV36747	PGI splice variant
36	329	17.1	90	20	AAV36746	PGI splice variant
37	328	17.0	374	22	AAV36745	Maize lysophosphat
38	326.5	17.0	374	21	AAV36744	Amino acid sequenc
39	326	16.9	374	15	AAV36743	Maize 2-acyltransf
40	317.5	16.5	374	21	AAV36742	Amino acid sequenc
41	311.5	16.2	387	21	AAV36741	Jojoba acyltransfe
42	310	16.1	377	17	AAV36740	Limnanthes CDNA cl
43	309.5	16.1	294	21	AAV36739	Zea mays protein f
44	308.5	16.0	389	21	AAV36738	Arabidopsis thalia
45	307.5	16.0	389	21	AAV36737	Arabidopsis thalia

ALIGNMENTS

RESULT 1	
AAV36729	standard; Protein; 364 AA.
ID	AAV36729;
AC	AAV36729;
DF	27-SEP-1999 (first entry)
DE	Human PGI protein sequence.
KW	PGI gene; biallelic marker; human; PSA; PGI-related biallelic marker;
KW	cancer; prostate cancer; diagnosis; therapy; prostate specific antigen.
OS	Homo sapiens.
PN	WO932644-A2.
PD	01-JUL-1999.
PF	22-DEC-1998; 98MO-IB02133.
PR	09-SEP-1998; 98US-0099658.
PR	22-DEC-1997; 97US-0096306.
XX	(GENSET ) GENSET.
XX	Blumenfeld M, Bougueleret L, Chumakov I, Cohen D;
XX	WPI: 1999-405178/34.
XX	N-PSDB; AA200870.
XX	Use of a prostate cancer associated gene and biallelic markers
XX	derived from it

PS Claim 7; Page 190-191; 385bp; English.

XX The invention relates to a mammalian PGI gene and protein, and a set of  
CC PGI allelic markers. The PGI polynucleotide and allelic markers are  
CC used in a hybridisation assay, a sequencing assay, or in an  
CC allele-specific amplification assay for determining the identity of a  
CC nucleotide at a PGI-related allelic marker. The methods can be used to  
CC detect and to assess the risk of developing cancer or prostate cancer.  
CC Early-stage diagnosis of prostate cancer relies on prostate specific  
CC antigen (PSA) dosage. However, the effectiveness of this is limited due  
CC to its inability to discriminate between malignant and non-malignant  
CC affections of the organ. A need exists for both a reliable diagnostic  
CC procedure which would enable early-stage diagnosis, and for preventative  
CC and curative treatments of the disease. The PGI gene can be used for  
CC detection of prostate cancer, and the risk of developing it in the  
CC future, and can also be used to determine therapies for the disease.

XX  
XX  
XX Sequence 364 AA:

QY 1 MLSLVLTHTYMKRYLPEVVLGTAPTYVLAAGVWRLLSATLPARFYALDRLKCYQS 60  
Db 1 MLSLVLTHTYMKRYLPEVVLGTAPTYVLAAGVWRLLSATLPARFYALDRLKCYQS 60  
QY 61 MVLFFENYTGVOILLYGDLDPKKNENTIIYLANHOSTVDWIVADILAIKRONLGHVRYLK 120  
Db 61 MVLFFENYTGVOILLYGDLDPKKNENTIIYLANHOSTVDWIVADILAIKRONLGHVRYLK 120  
QY 121 EGLMPLPGCYFQOHGSIYVKRSKAKFNEKEMRNKLSQSYVDAGTPMYLVIPPEGTRYNPE 180  
Db 121 EGLMPLPGCYFQOHGSIYVKRSKAKFNEKEMRNKLSQSYVDAGTPMYLVIPPEGTRYNPE 180  
QY 181 QTKVLSAQAAOAGLAVLKHVLTPTKATHVAFDCMKNLDAIYDVTYVEGKDSGQ 240  
Db 181 QTKVLSAQAAOAGLAVLKHVLTPTKATHVAFDCMKNLDAIYDVTYVEGKDSGQ 240  
QY 241 RRESPMTPEFLCKECPKIHIDRIDKKDVEDEHMRWHERPEIKDKMLIEFEYSPD 300  
Db 241 RRESPMTPEFLCKECPKIHIDRIDKKDVEDEHMRWHERPEIKDKMLIEFEYSPD 300  
QY 301 PERRRPEGKSVNSKLSIKTLPMSMLISGLTJAGMLMTDAGRKLVTWNTYVTLGCLMV 360  
Db 301 PERRRPEGKSVNSKLSIKTLPMSMLISGLTJAGMLMTDAGRKLVTWNTYVTLGCLMV 360  
QY 361 TTKA 364  
Db 361 TTKA 364

RESULT 2  
AAM50128  
ID AAM50128 standard; Protein: 364 AA.  
XX  
XX AAM50128:  
AC  
XX  
DT 21-Dec-2001 (first entry)  
XX  
DE Human acyltransferase 27417.  
XX  
XX Acyltransferase 27417: human; cell proliferation; cell migration;  
KM cell differentiation; ovary cancer; brain cancer; colon cancer;  
KM lung cancer; tumour; metastasis; sarcoma; carcinoma;  
KM adenocarcinoma; antitumour; lipid metabolism; diagnosis; therapy.  
XX  
XX Homo sapiens.  
OS  
XX  
FH Key Location/Qualifiers  
FT Peptide 1..49  
FT /label= Signal-peptide  
?

FT Protein 50..364  
FT /label= Mature\_protein  
FT Domain 50..320  
FT /label= N-terminal\_domain  
FT Domain 321..337  
FT /label= Transmembrane\_domain  
FT Domain 338..364  
FT /label= C-terminal\_domain  
FT Domain 71..363  
FT /label= Acyltransferase\_domain  
FT Modified-site 68..71  
FT /note= "Asn is N-glycosylated"  
FT Modified-site 241..244  
FT /note= "O-phosphorylated by cAMP- and cGMP-dependent protein kinase"  
FT Modified-site 11..13  
FT /note= "O-phosphorylated by protein kinase C"  
FT Modified-site 144..146  
FT /note= "O-phosphorylated by protein kinase C"  
FT Modified-site 205..207  
FT /note= "O-phosphorylated by protein kinase C"  
FT Modified-site 317..319  
FT /note= "O-phosphorylated by protein kinase C"  
FT Modified-site 361..363  
FT /note= "O-phosphorylated by protein kinase C"  
FT Modified-site 95..98  
FT /note= "O-phosphorylated by casein kinase II"  
FT Modified-site 158..161  
FT /note= "O-phosphorylated by casein kinase II"  
FT Modified-site 246..249  
FT /note= "O-phosphorylated by casein kinase II"  
FT Modified-site 82..89  
FT /note= "O-phosphorylated by tyrosine kinase"  
FT Modified-site 219..226  
FT /note= "O-phosphorylated by tyrosine kinase"  
FT Modified-site 23..28  
FT /note= "N-myristoylated"  
FT Modified-site 130..135  
FT /note= "N-myristoylated"  
FT Modified-site 330..335  
FT /note= "N-myristoylated"  
FT Modified-site 352..357  
FT /note= "N-myristoylated"  
FT Modified-site 230..343  
FT /note= "Amidation"  
FT Peptide 105..126  
FT /note= "predicted leucine zipper"  
XX  
XX W0200173051-A2.  
XX  
XX 04-OCT-2001.  
XX  
XX 26-MAR-2001; 2001WO-US09633.  
XX  
XX 24-MAR-2000; 2000US-192092P.  
XX  
XX (MILL-) MILLENNIUM PHARM INC.  
XX  
XX Meyers RA, Rudolph-Owen L, Macbeth KD, Tsai F, Williamson M;  
DR WPI; 2001-626267/72.  
XX  
XX N-PSDB; AAH27094.  
XX  
XX New polynucleotide encodes a polypeptide for the diagnosis and  
PT treatment of cancer and to alter cellular lipid metabolism comprises a  
PT polynucleotide encoding human acyltransferases designated 46743 and  
PT 27417 -  
XX  
XX  
XX Claim 9; Fig 8A-B; 150bp; English.  
PS  
XX  
XX The present sequence is that of a novel human protein, termed  
CC 27417, which shows the structural characteristics of members of the  
CC acyltransferase family, including an acyltransferase domain. The



CC invention provides novel acyltransferase 27417 nucleic acids and  
CC polypeptides, as well as methods for detecting their presence, and  
CC methods for screening for compounds that modulate their expression  
CC or activity. Such compounds can be used to treat conditions  
CC related to aberrant activity or expression of 27417 protein or  
CC nucleic acid, such as conditions involving deficient cellular  
CC proliferation, migration and/or differentiation. Also provided are  
CC methods for inhibiting the proliferation or migration, or inducing  
CC the killing, of a 27417-expressing cell, e.g. a hyperproliferative  
CC and/or metastatic cell, by contact with a compound that modulates  
CC the activity or expression of the 27417 protein or nucleic acid.  
CC The 27417-expressing cell is found in the healthy or diseased  
CC heart, blood vessels, kidney, skeletal muscle, brain or liver, or  
CC especially in a solid tumour, a soft tissue tumour or a metastatic  
CC lesion, a sarcoma, a carcinoma or an adenocarcinoma, and in  
CC particular in a hyperproliferative and/or metastatic cell found in  
CC ovarian, brain, colon or lung cancer. The compounds may also be  
CC used to modulate lipid metabolism in a 27417-expressing cell.  
CC Methods of disease diagnosis, e.g. by determining the  
CC presence of a genetic alteration in a 27417 polypeptide, and for  
CC evaluating the efficacy of a treatment of a disorder, are also  
CC provided.

XX Sequence 364 AA:

Query Match 100.0%: Score 1924; DB 22; Length 364;  
Best Local Similarity 100.0%: Pred. No. 5.8e-195;  
Matches 364; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 MLSTVLHTYSMRVLLPSVVLGTAPTVYLVANGVNRLLSAFLPARFYQALDDRLCYVOS 60  
DB 1 mlstvlhtysmrtyllpsvlllgtaptvylavgwvrllsaffparfyqalddrlycvygs 60  
OY 61 MYLFFEPENTGVQILLYGDLPRKNKENIYLANHGSTVDMIVADIIARONALGHRVYVK 120  
DB 61 mylffentygvqillygdlprkneniylanhgstvdmivadiiargnalghrvyvlk 120  
OY 121 EGLKMLPLTGYCYFAOHGIGYVRSKAFNEKEMRNKLOSVDAGTPMYLIVPEGRVYNE 180  
DB 121 eglkmlplygcyfaohgiglyvkrskafnekemrnklosvdaqtpmylviIpegrtynpe 180  
OY 181 QTKVLSASQAFAORGLAVLKHVLTPIKATVAFCMKNYDAIYDVYVVEGKDDGQ 240  
DB 181 qtkvlsasqafaarglavlkhlvtprikatvafcmknydaidydvvyvegkddgsq 240  
OY 241 RRESPTMTEFLCKECPKIHIDRIDKDVPEQEHMRRLHREPEIKDKMLIEFYESP 300  
DB 241 rresptmteflckeckpkihidiRIDKdvpeeqehmrtrwlherfeikdkmliefespd 300  
OY 301 PERRRRFPCKSVNSKLSIKKTLPSMLISGLFAGMLMTPDAGRKLVTNMTYIGTLGCLMV 360  
DB 301 perrrrfpcksvnsklsikktlpsmllsgltagmlmtdagrklyvntwlygtllgclmv 360  
OY 361 TIRKA 364  
DB 361 tika 364

RESULT 3

AAU00665 AAU00665 standard; Protein: 364 AA.

XX AC AAU00665;

XX DT 07-SEP-2001 (first entry)

DE Human lysophosphatidic acid acyltransferase isoform LPAAT-epsilon.

KW Lysophosphatidic acid acyltransferase epsilon; LPAAT-epsilon; LPA; PA;  
KW Lysophosphatidic acid; phosphatidic acid; acylation; cellular activation;  
KW phospholipid signalling; mitogenesis; inflammation; autoimmune disease;  
KW oncology; cancer; obesity; gene therapy.

XX Homo sapiens.  
OS  
XX  
XX WO200134782-A1.  
PN  
XX  
XX 17-MAY-2001.  
PD  
XX  
XX 02-NOV-2000; 2000WO-US30193.  
PF  
XX  
XX 09-NOV-1999; 99US-0436919.  
PR  
XX  
XX (CELL-) CELL THERAPEUTICS INC.  
PA  
XX  
XX Leung DW;  
PI  
XX  
XX WPI: 2001-335920/35.  
DR  
XX  
XX N-PSDB: AAS0649.

PT Novel isolated human isoform of lysophosphatidic acid  
PT acyltransferase-epsilon useful for diagnostic, therapeutic and  
PT screening purposes

Claim 1; Fig 1; 48pp; English.

The sequence represents a human lysophosphatidic acid acyltransferase  
(LPAAT) isoform, LPAAT-epsilon. LPAAT catalyses the acylation of  
CC lysophosphatidic acid (LPA) to phosphatidic acid (PA). LPA and PA have  
CC been identified as phospholipid signalling molecules that affect a wide  
CC range of biological responses. PA is involved in cellular activation and  
CC mitogenesis. Compounds that block PA generation and hence diminish lipid  
CC biosynthesis and the signal involved in cell activation are of  
CC therapeutic interest in the areas of inflammation and oncology (e.g.  
CC autoimmune diseases and cancer) as well as obesity treatment.  
CC LPAAT-epsilon and its corresponding DNA can be used in screening assays  
CC to detect agents that stimulate or inhibit the activity of LPAAT and,  
CC therefore, PA. The DNA is useful in tests to detect the presence or  
CC expression of LPAAT-epsilon in relation to certain diseases and  
CC conditions, and in disease prevention and treatment. The sequences of the  
CC invention are also useful for diagnosis of diseases and conditions in  
CC which the expression of LPAAT enzyme is abnormal.

XX Sequence 364 AA:

Query Match 100.0%: Score 1924; DB 22; Length 364;  
Best Local Similarity 100.0%: Pred. No. 5.8e-195;  
Matches 364; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 MLSTVLHTYSMRVLLPSVVLGTAPTVYLVANGVNRLLSAFLPARFYQALDDRLCYVOS 60  
DB 1 mlstvlhtysmrtyllpsvlllgtaptvylavgwvrllsaffparfyqalddrlycvygs 60  
OY 61 MYLFFEPENTGVQILLYGDLPRKNKENIYLANHGSTVDMIVADIIARONALGHRVYVK 120  
DB 61 mylffentygvqillygdlprkneniylanhgstvdmivadiiargnalghrvyvlk 120  
OY 121 EGLKMLPLTGYCYFAOHGIGYVRSKAFNEKEMRNKLOSVDAGTPMYLIVPEGRVYNE 180  
DB 121 eglkmlplygcyfaohgiglyvkrskafnekemrnklosvdaqtpmylviIpegrtynpe 180  
OY 181 QTKVLSASQAFAORGLAVLKHVLTPIKATVAFCMKNYDAIYDVYVVEGKDDGQ 240  
DB 181 qtkvlsasqafaarglavlkhlvtprikatvafcmknydaidydvvyvegkddgsq 240  
OY 241 RRESPTMTEFLCKECPKIHIDRIDKDVPEQEHMRRLHREPEIKDKMLIEFYESP 300  
DB 241 rresptmteflckeckpkihidiRIDKdvpeeqehmrtrwlherfeikdkmliefespd 300  
OY 301 PERRRRFPCKSVNSKLSIKKTLPSMLISGLFAGMLMTPDAGRKLVTNMTYIGTLGCLMV 360  
DB 301 perrrrfpcksvnsklsikktlpsmllsgltagmlmtdagrklyvntwlygtllgclmv 360  
OY 361 TIRKA 364

|||||  
Db 361 tika 364

## RESULT 4

AA041377  
ID AA041377 standard; protein: 372 AA.

XX AA041377:

DT 22-OCT-2001 (first entry)

XX Human polypeptide SEQ ID NO 6308.

XX Human: nootropic; immunosuppressant; cytostatic; gene therapy; cancer;  
KW peripheral nervous system; neuropathy; central nervous system; CNS;  
KW Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;  
KW amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;  
KW chemokine; thrombolytic; drug screening; arthritis; inflammation;  
KW leukemia.

XX Homo sapiens.

PN W0200153312-A1.

XX 26-JUL-2001.

PF 26-DEC-2000; 2000MO-US34263.

PR 21-JAN-2000; 2000US-0486725.

PR 25-APR-2000; 2000US-0552317.

PR 09-JUL-2000; 2000US-0598042.

PR 19-JUL-2000; 2000US-0620312.

PR 03-AUG-2000; 2000US-0653450.

PR 14-SEP-2000; 2000US-0662191.

PR 19-OCT-2000; 2000US-0693036.

PR 29-NOV-2000; 2000US-0727344.

XX (HYSE-) HYSEQ INC.

XX Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;  
PI Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J;  
PI Zhao QA, Zhou P, Goodrich R, Drmanac RT;

DR WPI: 2001-442253/47.

XX N-PSDB; AA160533.

XX Novel nucleic acids and polypeptides, useful for treating disorders  
PT such as central nervous system injuries -

PS Example 2; SEQ ID NO 6308; 10078pp; English.

XX The invention relates to human nucleic acids (AA157798-AA161369) and  
CC the encoded polypeptides (AA038642-AA042213) with nootropic,  
CC immunosuppressant and cytostatic activity. The polynucleotides are useful  
CC in gene therapy. A composition containing a polypeptide or polynucleotide  
CC of the invention may be used to treat diseases of the peripheral nervous  
CC system, such as peripheral nervous injuries, peripheral neuropathy and  
CC localized neuropathies and central nervous system diseases, such as  
CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic  
CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the  
CC utilisation of the activities such as: Immune system suppression,  
CC Activin/Inhibin activity, chemotactic/chemokinetic activity, haemostatic  
CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,  
CC assays for receptor activity, arthritis and inflammation, leukaemias and  
CC C.N.S disorders.  
CC Note: The sequence data for this patent did not form part of the printed  
CC specification.

XX Sequence 372 AA;

Query Match 99.4%; Score 1913; DB 22; Length 372;

Best Local Similarity 99.7%; Pred. No. 8.7e-194;  
Matches 363; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MLSTLVHTYSMRRLPSVVLGTAFTYVLAAGVWRLSAPLAPRFOALDDRLCYGOS 60

Db 9, mllstlvhtysmryllpsvvlgtapcyvlawgywrlsalfpftyqalddrlcyvys 68

QY 61 MVLFFENYTGVOILLYGDLPKKNENITLYANHOSTVDWIVADLAIQNALGHVRYLK 120

Db 69 mvlffenytygvqlllygdllpknkenillyanbqstvdwlvadllalqnalghvrylk 128

QY 121 EGLKMLPLVGCYFAOHGCIYVKSARKENKMKKLOSYVDAGTPMLVIFPBGTRNPE 180

Db 129 eglkmlplvygwyfaghygvyvksarknekmtrklqsyvdaagtpmlyvifpbgtrype 188

QY 181 QTKVLSSQKFAAORGAVLAKHVLTPRIKATHVAFDCMKVYLAIVYTVVYVYKDDGGQ 240

Db 189 qtkvlssqgaafagrgvavlkhvltprikatlvafdcmkxyldayvtyvvygkddggq 248

QY 241 RRESPTWTEFLCKECPKIHIDRIDKKDVPEDQEHMRMLHERFEIKDKMLIEFESPD 300

Db 249 rresptwteflckecpkihidridkkdvpeeqehmrmlherfekdkmliefespd 308

QY 301 PERKRRPGKSVNSKLSIKKTPSMLLSGLTAGMLMTDAGRKLYVNTWYVYGLGLVY 360

Db 309 perkrpfpgksvnsksikktpsmllsgltagmlmtdagrklyvntwylgcllvw 368

QY 361 TIKK 364

Db 369 tika 372

## RESULT 5

AA085532  
ID AA085532 standard; protein: 353 AA.

XX AA085532:

XX 25-SEP-2001 (first entry)

XX Human secreted protein (clone id HLT084).

XX Secreted protein; immunosuppressive; antiarthritic; antirheumatic;  
KW antiproliferative; cytostatic; cardiant; vasotropic; cerebroprotective;  
KW nootropic; neuroprotective; antibacterial; virucide; fungicide; human;  
KW opthalmological; gene therapy.

XX Homo sapiens.

XX W0200155430-A1.

XX 02-AUG-2001.

XX 17-JAN-2001; 2001MO-US01431.

XX 31-JAN-2000; 2000US-0179065.

XX 04-FEB-2000; 2000US-0180628.

XX 12-SEP-2000; 2000US-0231968.

XX (HUMA-) HUMAN GENOME SCI INC.

PI Rosen CA, Komitoulis GA, Baker KP, Birse CE, Soppet DR, Olsen HS;  
PI Moore PA, Wei P, Ebnier R, Duan DR, Shi Y, Choi GH, Fliscella M;  
PI Ni J, Ruben SM, Barash SC;

DR WPI: 2001-476220/51.

XX N-PSDB; AA046942.

XX 17 isolated nucleic acid molecules encoding human secreted proteins,  
PT used to preventing, treating or ameliorating a medical condition -

PS Claim 11; Page 450-451; 482pp; English.

CC The invention provides novel human secreted proteins and polynucleotides  
CC encoding them. The secreted proteins can be expressed by standard  
CC recombinant methodology. The secreted proteins and polynucleotides are  
CC used to prevent, treat or ameliorate a medical condition in e.g. humans,  
CC mice, rabbits, goats, horses, cats, dogs, chickens or sheep. They can  
CC also be used in diagnosing a pathological condition. The antibodies to  
CC the proteins can also be used in alleviating symptoms associated with the  
CC disorders and in diagnostic immunoassays e.g. radioimmunoassays or enzyme  
CC linked immunosorbent assays (ELISA). Disorders which are diagnosed or  
CC treated include autoimmune diseases e.g. rheumatoid arthritis,  
CC hyperproliferative disorders e.g. neoplasms of the breast or liver,  
CC cardiovascular disorders e.g. cardiac arrest, cerebrovascular disorders  
CC e.g. cerebral ischemia, angiodenesis, nervous system disorders e.g.  
CC Alzheimer's disease, infections caused by bacteria, viruses and fungi and  
CC ocular disorders e.g. corneal infection. The polypeptides can also be  
CC used to aid wound healing and epithelial cell proliferation, to prevent  
CC skin aging due to sunburn, to maintain organs before transplantation, for  
CC supporting cell culture of primary tissues, to regenerate tissues and in  
CC chemotaxis. The polypeptides can also be used as a food additive or  
CC preservative to increase or decrease storage capabilities. The present  
CC sequence represents a human secreted protein.

XX Sequence 353 AA:

Query Match 97.2%; Score 1871; DB 22; Length 353;  
Best Local Similarity 100.0%; Pred. No. 2.3e-189;  
Matches 353; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 12 MRYLPSVVLGTPAPYVYVAMGWRLISAFAPRPFYQALDDRLCYVQSMVLEFFENYTG 71  
DB 1 MYLLPSVVLIGTAPCYVLAWSYVWLLISAFIPARFYGALDDRLCYVQSMVLEFFENYTG 60  
QY 72 VOILLGDLPKNKENTIIYANHOSTVDWIVADILAIROMALGHVRYVLEKGLWLPYGC 131  
DB 61 VQILLYGDLPKXENIIYANHSTVDWIVADILAIROMALGHVRYVLEKGLWLPYGC 120  
QY 132 YFAOHGIIYVKSRAKFNEMERKLNQSYVDAGTPMYLVFPECTRYNPEQTKVLSAQAF 191  
DB 121 YFAGHGIIYVKSRAKFNEMERKLNQSYVDAGTPMYLVFPECTRYNPEQTKVLSAQAF 180  
QY 192 AAORGLAVLKHLVTPRIKATVAFDCMKNYLAIYDVTVVYEGKDDGGORRESPTMTLEFL 251  
DB 181 AAGRGLAVLKHLVTPRIKATVAFDCMKNYLAIYDVTVVYEGKDDGGORRESPTMTLEFL 240  
QY 252 CRECPRIHIIHIDRIKDDVPEDEHMRMLHEFEIKDKMLIFESPPDERKRRPGKS 311  
DB 241 CRECPRIHIIHIDRIKDDVPEDEHMRMLHEFEIKDKMLIFESPPDERKRRPGKS 300  
QY 312 VNSKLSIKKTLPSMLILSGLTAGMLMTDAGRKLVTWNTWYIGTLGLWWTIKA 364  
DB 301 VNSKLSIKKTLPSMLILSGLTAGMLMTDAGRKLVTWNTWYIGTLGLWWTIKA 353

RESULT 6

AAB85552  
ID AAB85552 standard; protein; 353 AA.

XX AAB85552;

DT 25-SEP-2001 (first entry)

DE Human secreted protein (clone id HSL1A81).

XX Secreted protein; immunosuppressive; antiarthritic; antirheumatic;  
KW antiproliferative; cytostatic; cardiatic; vasotrophic; cerebroprotective;  
KW nocitropic; neuroprotective; antibacterial; virucide; fungicide; human;  
XX opticalmallogical; gene therapy.

OS Homo sapiens.

XX WO200155430-A1.

XX

PD 02-AUG-2001.  
XX 17-JAN-2001; 2001WO-US01431.  
PF 31-JAN-2000; 2000US-0179065.  
PR 04-FEB-2000; 2000US-0180628.  
PR 12-SEP-2000; 2000US-0231968.  
XX (HUMA-) HUMAN GENOME SCI INC.  
PI Rosen CA, Komatsoulis GA, Baker KP, Birse CE, Soppet DR, Olsen HS;  
PI Moore PA, Wei P, Edner R, Duan DR, Shi Y, Choi GH, Fiscella M;  
PI Ni J, Ruben SM, Barash SC;  
XX WPI: 2001-476220/51.  
DR N-PSDB; AAM46962.  
PT 17 isolated nucleic acid molecules encoding human secreted proteins,  
PT used to preventing, treating or ameliorating a medical condition  
XX Claim 11; Page 465-466; 482pp; English.

CC The invention provides novel human secreted proteins and polynucleotides  
CC encoding them. The secreted proteins can be expressed by standard  
CC recombinant methodology. The secreted proteins and polynucleotides are  
CC used to prevent, treat or ameliorate a medical condition in e.g. humans,  
CC mice, rabbits, goats, horses, cats, dogs, chickens or sheep. They can  
CC also be used in diagnosing a pathological condition. The antibodies to  
CC the proteins can also be used in alleviating symptoms associated with the  
CC disorders and in diagnostic immunoassays e.g. radioimmunoassays or enzyme  
CC linked immunosorbent assays (ELISA). Disorders which are diagnosed or  
CC treated include autoimmune diseases e.g. rheumatoid arthritis,  
CC hyperproliferative disorders e.g. neoplasms of the breast or liver,  
CC cardiovascular disorders e.g. cardiac arrest, cerebrovascular disorders  
CC e.g. cerebral ischemia, angiodenesis, nervous system disorders e.g.  
CC Alzheimer's disease, infections caused by bacteria, viruses and fungi and  
CC ocular disorders e.g. corneal infection. The polypeptides can also be  
CC used to aid wound healing and epithelial cell proliferation, to prevent  
CC skin aging due to sunburn, to maintain organs before transplantation, for  
CC supporting cell culture of primary tissues, to regenerate tissues and in  
CC chemotaxis. The polypeptides can also be used as a food additive or  
CC preservative to increase or decrease storage capabilities. The present  
CC sequence represents a human secreted protein.

XX Sequence 353 AA:

Query Match 97.2%; Score 1871; DB 22; Length 353;  
Best Local Similarity 100.0%; Pred. No. 2.3e-189;  
Matches 353; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 12 MRYLPSVVLGTPAPYVYVAMGWRLISAFAPRPFYQALDDRLCYVQSMVLEFFENYTG 71  
DB 1 MYLLPSVVLIGTAPCYVLAWSYVWLLISAFIPARFYGALDDRLCYVQSMVLEFFENYTG 60  
QY 72 VOILLGDLPKNKENTIIYANHOSTVDWIVADILAIROMALGHVRYVLEKGLWLPYGC 131  
DB 61 VQILLYGDLPKXENIIYANHSTVDWIVADILAIROMALGHVRYVLEKGLWLPYGC 120  
QY 132 YFAOHGIIYVKSRAKFNEMERKLNQSYVDAGTPMYLVFPECTRYNPEQTKVLSAQAF 191  
DB 121 YFAGHGIIYVKSRAKFNEMERKLNQSYVDAGTPMYLVFPECTRYNPEQTKVLSAQAF 180  
QY 192 AAORGLAVLKHLVTPRIKATVAFDCMKNYLAIYDVTVVYEGKDDGGORRESPTMTLEFL 251  
DB 181 AAGRGLAVLKHLVTPRIKATVAFDCMKNYLAIYDVTVVYEGKDDGGORRESPTMTLEFL 240  
QY 252 CRECPRIHIIHIDRIKDDVPEDEHMRMLHEFEIKDKMLIFESPPDERKRRPGKS 311  
DB 241 CRECPRIHIIHIDRIKDDVPEDEHMRMLHEFEIKDKMLIFESPPDERKRRPGKS 300  
QY 312 VNSKLSIKKTLPSMLILSGLTAGMLMTDAGRKLVTWNTWYIGTLGLWWTIKA 364  
DB 301 VNSKLSIKKTLPSMLILSGLTAGMLMTDAGRKLVTWNTWYIGTLGLWWTIKA 353

Db 301 vnsklsikktlpsmlllsgltagmlmtcdagrklyvntwlygcllgclwltlka 353

RESULT 7

ID AAB93595 standard; Protein; 353 AA.

XX AAB93595;

XX 26-JUN-2001 (first entry)

XX Human protein sequence SEQ ID NO:13028.

XX Human; primer; detection; diagnosis; antisense therapy; gene therapy.

XX Homo sapiens.

XX EP1074617-A2.

XX 07-FEB-2001.

XX 28-JUL-2000; 2000EP-0116126.

XX 29-JUL-1999; 99JP-0248036.

XX 27-AUG-1999; 99JP-0300253.

XX 11-JAN-2000; 2000JP-0116776.

XX 02-MAY-2000; 2000JP-0183767.

XX 09-JUN-2000; 2000JP-0241899.

XX (HELI-) HELIX RES INST.

XX Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;

XX Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;

XX WPI; 2001-318749/34.

XX PT Primer sets for synthesizing polynucleotides, particularly the 5602

XX full-length cDNAs defined in the specification, and for the detection

XX and/or diagnosis of the abnormality of the proteins encoded by the

XX full-length cDNAs -

XX PT full-length cDNAs -

XX Claim 8; SEQ ID 13028; 2537pp + CD ROM; English.

XX The present invention describes primer sets for synthesizing 5602

XX full-length cDNAs defined in the specification. Where a primer set

XX comprises: (a) an oligo-dT primer and an oligonucleotide complementary

XX to the complementary strand of a polynucleotide which comprises one of

XX the 5602 nucleotide sequences defined in the specification, where the

XX oligonucleotide comprises at least 15 nucleotides; or (b) a combination

XX of an oligonucleotide comprising a sequence complementary to the

XX complementary strand of a polynucleotide which comprises a 5'-end

XX sequence and an oligonucleotide comprising a sequence complementary to a

XX polynucleotide which comprises a 3'-end sequence, where the

XX oligonucleotide comprises at least 15 nucleotides and the combination of

XX the 5'-end sequence/3'-end sequence is selected from those defined in

XX the specification. The primer sets can be used in antisense therapy and

XX in gene therapy. The primers are useful for synthesizing polynucleotides,

XX particularly full-length cDNAs. The primers are also useful for the

XX detection and/or diagnosis of the abnormality of the proteins encoded by

XX the full-length cDNAs. The primers allow obtaining of the full-length

XX cDNAs easily without any specialised methods. AAH03166 to AAH13628 and

XX AAH13633 to AAH18742 represent human cDNA sequences; AAB92446 to

XX AAB95893 represent human amino acid sequences; and AAH13629 to AAH13632

XX represent oligonucleotides, all of which are used in the exemplification

XX of the present invention.

XX Sequence 353 AA;

Query Match 97.1%; Score 1868; DB 22; Length 353;

Best Local Similarity 99.7%; Pred. No. 4.7e-189;

Matches 352; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 12 MRYLPSVVLGTAPTYVLAMGWRLLSAPLPARFYQALDDRLCYQSVNLEFFENYTG 71

Db 1 mrylpsvvlgtaptcyvllawgvrllsalfparfyqalddrlycvgsnmvlffennytg 60

QY 72 VOILLYGDLPKRNENIYYLANHOSTVDMIVADILAIRONALGHVRVYLKGGKWLPLYGC 131

Db 61 vqilllygdlpknenillylanhgstvwivadiatlrqnalghvrlylkeglwlplygc 120

QY 132 YFAOHGCIYVKRSARFENKEKRNKLOSVDAGTPMYLVIFPECTRYNPEOTKVLASQAQF 191

Db 121 yfaohggllyvkrsakfkekenrkvgsyvdagtpmylvifpegrtrynpeotkvlasaqaf 180

QY 192 AAGRCGLAVLKHVLTTPRIKATHVAFDCKKNYLDALIVDTVYVEGKDDGQGRRESPTMEFL 251

Db 181 aagrglavlkhlvtprikathtvafdcnknyldaldvtvvyegkdqgqrresptmefl 240

QY 252 CKECPKIHINIDRTDKDVPDEQHMRRWLHERREIDKMLIEFESPDPERRRRPCKS 311

Db 241 ckecpklnhldrtddkdvpeeqhmrrwlherfelkdkmliefespdpertrrpkgs 300

QY 312 VNSKLSIKKTLPMSMLITSGLTAGMLMTDAGRKLTVNNTWYGTLLGCLWVLIKA 364

Db 301 vnsklsikktlpsmlllsgltagmlmtcdagrklyvntwlygcllgclwltlka 353

RESULT 8

AAV36728

ID AAV36728 standard; Protein; 352 AA.

XX AAV36728;

XX 27-SEP-1999 (first entry)

XX Human PGI protein sequence.

XX PGI gene; biallelic marker; human; PSA; PGI-related biallelic marker;

XX cancer; prostate cancer; diagnosis; therapy; prostate specific antigen.

XX Homo sapiens.

XX WO9932644-A2.

XX 01-JUL-1999.

XX 22-DEC-1998; 98WO-1B02133.

XX 09-SEP-1998; 98US-0099658.

XX 22-DEC-1997; 97US-0996306.

XX (GBST) GENSET.

XX Blumenfeld M, Bougueleret L, Chumakov I, Cohen D;

XX WPI; 1999-405178/34.

XX N-PSDB; AAZ00872.

XX use of a prostate cancer associated gene and biallelic markers

XX derived from it

XX Claim 7; Page 190; 385pp; English.

XX The invention relates to a mammalian PGI gene and protein, and a set of

XX PGI biallelic markers. The PGI polynucleotide and biallelic markers are

XX used in a hybridisation assay, a sequencing assay, or in an

XX allele-specific amplification assay for determining the identity of a

XX nucleotide at a PGI-related biallelic marker. The methods can be used to

XX detect and to assess the risk of developing cancer or prostate cancer.

XX Early-stage diagnosis of prostate cancer relies on prostate specific

XX antigen (PSA) dosage. However, the effectiveness of this is limited due

XX to its inability to discriminate between malignant and non-malignant

XX affections of the organ. A need exists for both a reliable diagnostic

XX procedure which would enable early-stage diagnosis, and for preventative

XX and curative treatments of the disease. The PGI gene can be used for

PI Zhao QA, Zhou P, Goodrich R, Drmanac RT;  
XX  
DR WPI; 2001-442253/47.  
DR N-PSDB; AA158747.

PT such as central nervous system injuries -  
XX  
PS Example 4; SEQ ID NO 2736; 10078bp; English.

CC The invention relates to human nucleic acids (AA157798-AA161369) and  
CC the encoded polypeptides (AAM38642-AAM42213) with neurotrophic,  
CC immunosuppressant and cytostatic activity. The polynucleotides are useful  
CC in gene therapy. A composition containing a polypeptide or polynucleotide  
CC of the invention may be used to treat diseases of the peripheral nervous  
CC system, such as peripheral nervous injuries, peripheral neuropathy and  
CC localised neuropathies and central nervous system diseases, such as  
CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic  
CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the  
CC utilisation of the activities such as: Immune system suppression,  
CC Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic  
CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,  
CC assays for receptor activity, arthritis and inflammation, leukaemias and  
CC C.N.S disorders.  
CC Note: The sequence data for this patent did not form part of the printed  
CC specification.  
CX  
SQ Sequence 353 AA;

QY	12	MRLLEPVLLGTPAPRYVLAMGVWRLLSFLPAREVQLADDRICVVYSMWLPFEENVTG	71
Dd	1	mryllpsvvllygtapcyvlaagwvrllsaflparfygaiddrlcvcygsmylfftenytg	60
QY	72	VOILLTEDELRKNEENTIIYLANHSTVDWIIVADIILAIRONALGHARYLKEGLMKMLPLYGC	131
Dd	61	vqllllvgdlprkxkenlllylanhqstcvdrvalailatrqnalghrvyylkeglkwlpdygc	120
QY	132	YFAOHGGIYVKRSAKFNEKEEMRNKLQSYVDAGTPMYLVIPEBGRTRINDEQTIVLASQAQF	191
Dd	121	yfaehggilykkrzakfnekemnrnklsqsydaqctpmylvlfpegrtynpedklvtsasqaf	180
QY	192	AAGRGALAVLKHNVLTPRIKATHTVAFDCKMKNYLDAYIDVLYVVEKGDDGQRREESTMTMEFL	251
Dd	181	aagrgalavlkhnvltprikathvafdcmknyla ydvlvvyevgkdgdqqrresptmtefl	240
QY	252	CKECPKHINHIDRIDKKDVP EEGEHHRRMLAHREFIKDKMLEIEFSHPDERRRRFPFGKS	311
Dd	241	ckecpkhinhidridkkdvp eegehmrrvlnherfeikdkmliefeshpderrrrpfpgks	300
QY	312	VNSKLSIKKTLPSPMLILSGITAGMLMTDAGRKYLVNTWVIYGTLGCLMTWTIKA	364
Dd	301	vnsklsikktpspmlilsgitagmlmtdagrklylvntwiygclwtika	353
 RESULT 10 AAAY36751 ID AAY36751 standard; Protein; 315 AA. XX AC AAY36751; DT DT 27-SEP-1999 (first entry) XX XN PC1 splice variant allele protein sequence. XX KM PC1 gene; biallelic marker; human; PSA; PGI-related biallelic marker; XX KM cancer; prostate cancer; diagnosis; therapy; prostate specific antigen XS Homo sapiens.			









PR 08-NOV-2000; 2000US-0246478.  
PR 08-NOV-2000; 2000US-0246523.  
PR 08-NOV-2000; 2000US-0246524.  
PR 08-NOV-2000; 2000US-0246525.  
PR 08-NOV-2000; 2000US-0246526.  
PR 08-NOV-2000; 2000US-0246527.  
PR 08-NOV-2000; 2000US-0246528.  
PR 08-NOV-2000; 2000US-0246532.  
PR 08-NOV-2000; 2000US-0246609.  
PR 08-NOV-2000; 2000US-0246610.  
PR 08-NOV-2000; 2000US-0246611.  
PR 08-NOV-2000; 2000US-0246613.  
PR 17-NOV-2000; 2000US-0249207.  
PR 17-NOV-2000; 2000US-0249208.  
PR 17-NOV-2000; 2000US-0249209.  
PR 17-NOV-2000; 2000US-0249210.  
PR 17-NOV-2000; 2000US-0249211.  
PR 17-NOV-2000; 2000US-0249212.  
PR 17-NOV-2000; 2000US-0249213.  
PR 17-NOV-2000; 2000US-0249214.  
PR 17-NOV-2000; 2000US-0249215.  
PR 17-NOV-2000; 2000US-0249216.  
PR 17-NOV-2000; 2000US-0249217.  
PR 17-NOV-2000; 2000US-0249218.  
PR 17-NOV-2000; 2000US-0249244.  
PR 17-NOV-2000; 2000US-0249245.  
PR 17-NOV-2000; 2000US-0249264.  
PR 17-NOV-2000; 2000US-0249265.  
PR 17-NOV-2000; 2000US-0249297.  
PR 17-NOV-2000; 2000US-0249299.  
PR 17-NOV-2000; 2000US-0249300.  
PR 01-DEC-2000; 2000US-0250160.  
PR 01-DEC-2000; 2000US-0250391.  
PR 05-DEC-2000; 2000US-0251030.  
PR 05-DEC-2000; 2000US-0251988.  
PR 05-DEC-2000; 2000US-0256719.  
PR 06-DEC-2000; 2000US-0251479.  
PR 08-DEC-2000; 2000US-0251856.  
PR 08-DEC-2000; 2000US-0251868.  
PR 08-DEC-2000; 2000US-0251869.  
PR 08-DEC-2000; 2000US-0251989.  
PR 08-DEC-2000; 2000US-0251990.  
PR 11-DEC-2000; 2000US-0254097.  
PR 05-JAN-2001; 2001US-02559678.  
XX  
XX  
PA (HUMA-) HUMAN GENOME SCI INC.  
XX  
PI Rosen CA, Barash SC, Ruben SM;  
XX  
XX WPI: 2001-457716/49.  
DR N-PSDB; AAD16620.  
XX  
XX  
PT Isolated lipid metabolism polypeptide for screening to identify  
PT antagonists and agonists that may enhance or block activities mediated  
PT by lipid metabolism proteins and also for testing and detection e.g.  
PT diagnosis -  
XX  
XX  
PS Claim 11; SEQ ID NO 47; 453bp; English.  
XX  
XX  
CC AAD16606-AAD16627 represent cDNAs corresponding to human lipid metabolism  
CC protein genes, and AAE09639-AAE09660 represent the proteins they encode.  
CC AAD16628-AAD16647 represent human lipid metabolism genomic DNAs. The  
CC genes and their corresponding proteins are useful for preventing,  
CC treating or ameliorating medical conditions such as immune disorders,  
CC inflammatory disorders (e.g., conjunctivitis, glossitis), autoimmune  
CC diseases (e.g., rheumatoid arthritis, systemic lupus erythematosus),  
CC respiratory disorders (asthma, allergy), CNS disorders (e.g., Alzheimer's  
CC disease, Parkinson's disease), haematopoietic disorders, cardiovascular  
CC disorders (e.g., atherosclerosis, coronary disease), hyperlipidaemias,  
CC hypolipidaemia, lipidoses, familial hypercholesterolaemia, Gaucher's  
CC disease, Tay-Sachs's disease, diabetes mellitus, cancer and other enzyme  
CC abnormalities. They are used for the prevention of mental retardation  
CC and the enlargement of spleen and liver. They are used in gene therapy

CC and antisense therapy. The present sequence represents a human lipid  
CC metabolism protein of the invention.  
CC Note: The sequence data for this patent did not form part of the printed  
CC specification, but was obtained in electronic format directly from WIPO  
CC at ftp.wipo.int/pub/published\_pct\_sequences.  
XX  
XX  
SQ Sequence 269 AA:  
  
Query Match 74.2%; Score 1427; DB 22; Length 269;  
Best Local Similarity 100.0%; Pred. No. 1.7e-142;  
Matches 269; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 96 TVDMIVADIIAIRQNALGHRVYLKGLKWLPLYGCFPAOHGGIYKRSAPFNEKEKRNK 155  
DB 1 tvdwivadiiairgnalghrvylkeglwpjlygcyfaaghylykrsakfnekemrnk 60  
  
QY 156 LQSYVDAGTPMYLVIPEEGTRYNPEQTKVLSASQAPAAORGAVLNRHVLTPRIKATNVAP 215  
DB 1 lqsyvdagtpmylvifpegtrynpeqtkvlssaqaafaagrlavlkhvltprikachvaf 120  
  
QY 216 DCKKNLDAIYDVTVYVEGKDDGGRRRESPTMEFLCKECPKHIIHDIRDKDVEEEOE 275  
DB 121 dcmknyldaiydvltvyegkddgqrrresptmeflickecpkhihdiridkdvpeeqe 180  
  
QY 276 HMRRLWHERPEIKDKMLIEFRESPDERRRKRFPGKSVNSKLSIKTLPMSLILSGLTAGM 335  
DB 181 hmrrlwherfelkdmliefrespderrkrfpgksvnskslsiktlpsmlisgltagm 240  
  
QY 336 LMTDAGRKLVTNTWIVGTLLGCLMWTIKA 364  
DB 241 lmtdagrklvntwivgtllgclwtika 269  
  
RESULT 15  
AAY36745  
ID AAY36745 standard; Protein; 261 AA.  
XX  
AC AAY36745;  
XX  
DT 27-SEP-1999 (first entry)  
XX  
DE PGI splice variant allele protein sequence.  
XX  
KW PGI gene: biallelic marker; human; PSA: PGI-related biallelic marker;  
KW cancer; prostate cancer; diagnosis; therapy; prostate specific antigen.  
XX  
OS Homo sapiens.  
OS  
PN WO9332644-A2.  
PN  
XX  
PD 01-JUL-1999.  
XX  
XX  
PF 22-DEC-1998; 98WO-IB02133.  
XX  
PR 09-SEP-1998; 98US-0099658.  
XX  
PR 22-DEC-1997; 97US-0996306.  
XX  
PA (GEST ) GENSET.  
XX  
PI Blumenfeld M, Bougueleret L, Chumakov I, Cohen D;  
XX  
XX WPI: 1999-405178/34.  
DR N-PSDB; AA200971.  
XX  
XX  
PT Use of a prostate cancer associated gene and biallelic markers  
PT derived from it  
XX  
XX  
PS Claim 7; Page 251; 385bp; English.  
XX  
XX  
CC The invention relates to a mammalian PGI gene and protein, and a set of  
CC PGI biallelic markers. The PGI polynucleotide and biallelic markers are  
CC used in a hybridisation assay, a sequencing assay, or in an



GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 28, 2002, 11:08:10 ; Search time 31 Seconds  
(without alignments)  
286.804 Million cell updates/sec

Title: US-09-853-526-5  
Perfect score: 1924  
Sequence: 1 MLTSLVLTYSMRYLPSVY.....YVNTWIVGTLLGLMTWTKA 364

Scoring table:  
BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 231628 seqs, 24425594 residues

Total number of hits satisfying chosen parameters: 231628

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-Processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :  
Issued\_Patents\_AA:\*  
1: /cgn2\_6/prodata/2/1aa/5A.COMB.pep:\*  
2: /cgn2\_6/prodata/2/1aa/5B.COMB.pep:\*  
3: /cgn2\_6/prodata/2/1aa/6A.COMB.pep:\*  
4: /cgn2\_6/prodata/2/1aa/6B.COMB.pep:\*  
5: /cgn2\_6/prodata/2/1aa/CTUS.COMB.pep:\*  
6: /cgn2\_6/prodata/2/1aa/Backfilest1.pep:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	length	DB	ID	Description
1	1924	100.0	364	2	US-08-996-306-5	Sequence 5, Appl
2	1924	100.0	364	4	US-09-338-907-5	Sequence 5, Appl
3	1924	100.0	364	4	US-09-218-207-5	Sequence 5, Appl
4	1871	97.2	353	2	US-08-996-306-4	Sequence 4, Appl
5	1871	97.2	353	4	US-09-338-907-4	Sequence 4, Appl
6	1871	97.2	353	4	US-09-218-207-4	Sequence 4, Appl
7	1644	85.4	315	4	US-09-338-907-134	Sequence 134, App
8	1644	85.4	315	4	US-09-218-207-134	Sequence 134, App
9	1568.5	81.5	354	4	US-09-338-907-74	Sequence 74, Appl
10	1568.5	81.5	354	4	US-09-218-207-74	Sequence 74, Appl
11	1555.5	80.8	300	4	US-09-338-907-135	Sequence 135, App
12	1555.5	80.8	300	4	US-09-218-207-135	Sequence 135, App
13	1498	77.9	291	4	US-09-338-907-127	Sequence 127, App
14	1498	77.9	291	4	US-09-218-207-127	Sequence 127, App
15	1335	68.9	261	4	US-09-338-907-136	Sequence 128, App
16	1335	68.9	261	4	US-09-218-207-136	Sequence 128, App
17	1203	62.5	228	4	US-09-338-907-70	Sequence 70, Appl
18	1203	62.5	228	4	US-09-218-207-70	Sequence 70, Appl
19	1182.5	61.5	238	4	US-09-338-907-126	Sequence 126, App
20	1182.5	61.5	238	4	US-09-218-207-126	Sequence 126, App
21	976	50.7	185	4	US-09-338-907-136	Sequence 136, App
22	976	50.7	185	4	US-09-218-207-136	Sequence 136, App
23	820	42.6	182	4	US-09-338-907-133	Sequence 133, App
24	820	42.6	182	4	US-09-218-207-133	Sequence 133, App
25	449	23.3	97	4	US-09-338-907-132	Sequence 132, App
26	449	23.3	97	4	US-09-218-207-132	Sequence 132, App
27	332	17.3	77	4	US-09-338-907-125	Sequence 125, App

28	332	17.3	77	4	US-09-218-207-125	Sequence 125, App
29	330	17.2	66	4	US-09-338-907-131	Sequence 131, App
30	330	17.2	66	4	US-09-218-207-131	Sequence 131, App
31	329	17.1	68	4	US-09-338-907-130	Sequence 130, App
32	329	17.1	68	4	US-09-218-207-130	Sequence 130, App
33	329	17.1	90	4	US-09-338-907-129	Sequence 129, App
34	329	17.1	90	4	US-09-218-207-129	Sequence 129, App
35	328	17.0	374	2	US-08-454-267-2	Sequence 2, Appl
36	328	17.0	374	2	US-08-941-319-2	Sequence 2, Appl
37	328	17.0	374	2	US-08-941-319-6	Sequence 6, Appl
38	328	17.0	374	4	US-09-035-098-2	Sequence 2, Appl
39	328	17.0	374	4	US-09-035-098-6	Sequence 6, Appl
40	328	17.0	374	4	US-09-215-252-5	Sequence 5, Appl
41	328	17.0	374	4	US-09-215-252-17	Sequence 17, Appl
42	268.5	14.0	376	4	US-09-215-252-13	Sequence 13, Appl
43	240	12.5	295	2	US-08-454-267-7	Sequence 7, Appl
44	239	12.4	295	2	US-08-941-319-7	Sequence 7, Appl
45	239	12.4	295	2	US-08-941-319-7	Sequence 7, Appl

## ALIGNMENTS

RESULT 1  
US-08-996-306-5  
; Sequence 5, Application US/08996306  
; Patent No. 5945522  
; GENERAL INFORMATION:  
; APPLICANT: Cohen, Daniel  
; APPLICANT: Chumakov, Ilya  
; APPLICANT: Blumenfeld, Marta  
; APPLICANT: Bougenleret, Lydie  
; TITLE OF INVENTION: Prostate cancer gene  
; NUMBER OF SEQUENCES: 68  
; CORRESPONDENCE ADDRESSES:  
; ADDRESSEE: Knobbe, Martens, Olson & Bear  
; STREET: 501 West Broadway  
; CITY: San Diego  
; STATE: California  
; COUNTRY: USA  
; ZIP: 92101-3505  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy Disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: Win95  
; SOFTWARE: Word  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/996,306  
; FILING DATE:  
; CLASSIFICATION:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Israelson, Ned A.  
; REGISTRATION NUMBER: 29,655  
; REFERENCE/DOCKET NUMBER: GENSET.018A  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (619) 235-8550  
; TELEFAX: (619) 235-0176  
; INFORMATION FOR SEQ ID NO: 5:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 364 amino acids  
; TYPE: AMINO ACID  
; STRANDEDNESS: SINGLE  
; TOPOLOGY: LINEAR  
; MOLECULE TYPE: PROTEIN  
; ORIGINAL SOURCE:  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: potential protein  
; LOCATION: 1..364  
US-08-996-306-5  
Query Match 100.0%; Score 1924; DB 2; Length 364;

```
Best Local Similarity 100.0%; Pred. No. 4.2e-200;
Matches 364; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MLTSLVHTTYSMRYLLPSVVLGTAPTYVLAMGYWRLSAFLPARFYQALDRLCYCYOS 60
Db 1 MLTSLVHTTYSMRYLLPSVVLGTAPTYVLAMGYWRLSAFLPARFYQALDRLCYCYOS 60
QY 61 MVLFFENYTGVOILLGDLDPKKNENITYLANHSTYDVIADILAIQNALGHVRYLAK 120
Db 61 MVLFFENYTGVOILLGDLDPKKNENITYLANHSTYDVIADILAIQNALGHVRYLAK 120
QY 121 EGLKWLPLGYCYFAOHGIGIYKRSKAKFNEKEMRNKLOSVDAGTPMYLVIFPESTRNPE 180
Db 121 EGLKWLPLGYCYFAOHGIGIYKRSKAKFNEKEMRNKLOSVDAGTPMYLVIFPESTRNPE 180
QY 181 QTKVLSASQAFAAOAGLAVLKHVLTLPRIKATHVAFDCMKNYLDIYDVTVYEGKDDGQ 240
Db 181 QTKVLSASQAFAAOAGLAVLKHVLTLPRIKATHVAFDCMKNYLDIYDVTVYEGKDDGQ 240
QY 241 RRESPTMTEFLCKECPKIHIDRIDKKDVPEDQEHMRWLHERFEIKDKMLIEFYESPD 300
Db 241 RRESPTMTEFLCKECPKIHIDRIDKKDVPEDQEHMRWLHERFEIKDKMLIEFYESPD 300
QY 301 PERKRRPFGKSVNSKLSIKTLPMSMLISGLTAGMLMTDAGRKLIVNTWITGTLGCLMW 360
Db 301 PERKRRPFGKSVNSKLSIKTLPMSMLISGLTAGMLMTDAGRKLIVNTWITGTLGCLMW 360
QY 361 TTKA 364
Db 361 TTKA 364

RESULT 2
US-09-338-907-5
; Sequence 5, Application US/09338907
; Patent No. 6265546
; GENERAL INFORMATION:
; APPLICANT: Cohen, Daniel
; APPLICANT: Blumenfeld, Marta
; APPLICANT: Ilya, Chumakov
; APPLICANT: Bougueleret, Lydie
; TITLE OF INVENTION: PROSTATE CANCER GENE
; FILE REFERENCE: GENSET.18C11CP
; CURRENT APPLICATION NUMBER: US/09/338, 907
; EARLIER FILING DATE: 1999-06-23
; EARLIER APPLICATION NUMBER: 08/996,306
; EARLIER FILING DATE: 1997-12-22
; EARLIER APPLICATION NUMBER: 09/099,658
; EARLIER FILING DATE: 1998-09-09
; EARLIER APPLICATION NUMBER: 09/218,207
; NUMBER OF SEQ ID NOS: 578
; SOFTWARE: Patent.pm
; SEQ ID NO 5
; LENGTH: 364
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-338-907-5

Query Match 100.0%; Score 1924; DB 4; Length 364;
Best Local Similarity 100.0%; Pred. No. 4.2e-200;
Matches 364; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Db 121 EGLKWLPLGYCYFAOHGIGIYKRSKAKFNEKEMRNKLOSVDAGTPMYLVIFPESTRNPE 180
QY 181 QTKVLSASQAFAAOAGLAVLKHVLTLPRIKATHVAFDCMKNYLDIYDVTVYEGKDDGQ 240
Db 181 QTKVLSASQAFAAOAGLAVLKHVLTLPRIKATHVAFDCMKNYLDIYDVTVYEGKDDGQ 240
QY 241 RRESPTMTEFLCKECPKIHIDRIDKKDVPEDQEHMRWLHERFEIKDKMLIEFYESPD 300
Db 241 RRESPTMTEFLCKECPKIHIDRIDKKDVPEDQEHMRWLHERFEIKDKMLIEFYESPD 300
QY 301 PERKRRPFGKSVNSKLSIKTLPMSMLISGLTAGMLMTDAGRKLIVNTWITGTLGCLMW 360
Db 301 PERKRRPFGKSVNSKLSIKTLPMSMLISGLTAGMLMTDAGRKLIVNTWITGTLGCLMW 360
QY 361 TTKA 364
Db 361 TTKA 364

RESULT 3
US-09-218-207-5
; Sequence 5, Application US/09218207
; Patent No. 6346381
; GENERAL INFORMATION:
; APPLICANT: Cohen, Daniel
; APPLICANT: Blumenfeld, Marta
; APPLICANT: Ilya, Chumakov
; APPLICANT: Bougueleret, Lydie
; TITLE OF INVENTION: Prostate cancer gene
; FILE REFERENCE: GENSET.018CP1
; CURRENT APPLICATION NUMBER: US/09/218,207
; EARLIER FILING DATE: 1998-12-22
; EARLIER APPLICATION NUMBER: 08/996,306
; EARLIER FILING DATE: 1997-12-22
; EARLIER APPLICATION NUMBER: 60/099,658
; EARLIER FILING DATE: 1998-09-09
; NUMBER OF SEQ ID NOS: 578
; SOFTWARE: Patent.pm
; SEQ ID NO 5
; LENGTH: 364
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-218-207-5

Query Match 100.0%; Score 1924; DB 4; Length 364;
Best Local Similarity 100.0%; Pred. No. 4.2e-200;
Matches 364; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MLTSLVHTTYSMRYLLPSVVLGTAPTYVLAMGYWRLSAFLPARFYQALDRLCYCYOS 60
Db 1 MLTSLVHTTYSMRYLLPSVVLGTAPTYVLAMGYWRLSAFLPARFYQALDRLCYCYOS 60
QY 61 MVLFFENYTGVOILLGDLDPKKNENITYLANHSTYDVIADILAIQNALGHVRYLAK 120
Db 61 MVLFFENYTGVOILLGDLDPKKNENITYLANHSTYDVIADILAIQNALGHVRYLAK 120
QY 121 EGLKWLPLGYCYFAOHGIGIYKRSKAKFNEKEMRNKLOSVDAGTPMYLVIFPESTRNPE 180
Db 121 EGLKWLPLGYCYFAOHGIGIYKRSKAKFNEKEMRNKLOSVDAGTPMYLVIFPESTRNPE 180
QY 181 QTKVLSASQAFAAOAGLAVLKHVLTLPRIKATHVAFDCMKNYLDIYDVTVYEGKDDGQ 240
Db 181 QTKVLSASQAFAAOAGLAVLKHVLTLPRIKATHVAFDCMKNYLDIYDVTVYEGKDDGQ 240
QY 241 RRESPTMTEFLCKECPKIHIDRIDKKDVPEDQEHMRWLHERFEIKDKMLIEFYESPD 300
Db 241 RRESPTMTEFLCKECPKIHIDRIDKKDVPEDQEHMRWLHERFEIKDKMLIEFYESPD 300
QY 301 PERKRRPFGKSVNSKLSIKTLPMSMLISGLTAGMLMTDAGRKLIVNTWITGTLGCLMW 360
Db 301 PERKRRPFGKSVNSKLSIKTLPMSMLISGLTAGMLMTDAGRKLIVNTWITGTLGCLMW 360
```

OY 361 TIKa 364  
Db 361 TIKa 364

RESULT 4  
US-08-996-306-4  
Sequence 4, Application US/08996306  
Patent No. 5945522  
GENERAL INFORMATION:  
APPLICANT: Cohen, Daniel  
APPLICANT: Chumakov, Ilya  
APPLICANT: Blumenfeld, Marta  
APPLICANT: Bougueleret, Lydie  
TITLE OF INVENTION: Prostate cancer gene  
NUMBER OF SEQUENCES: 68  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Knobbe, Martens, Olson & Bear  
STREET: 501 West Broadway  
CITY: San Diego  
STATE: California  
COUNTRY: USA  
ZIP: 92101-3505  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy Disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: Win95  
SOFTWARE: Word  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/996,306  
FILING DATE:  
CLASSIFICATION:  
ATTORNEY/AGENT INFORMATION:  
NAME: Israelsen, Ned A.  
REGISTRATION NUMBER: 29,655  
REFERENCE/DOCKET NUMBER: GENSET.018A  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (619) 235-8550  
TELEFAX: (619) 235-0176  
INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 353 amino acids  
TYPE: AMINO ACID  
STRANDEDNESS: SINGLE  
TOPOLOGY: LINEAR  
MOLECULE TYPE: PROTEIN  
ORIGINAL SOURCE:  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: potential Transmembrane helix  
LOCATION: 1..33  
IDENTIFICATION METHOD: Rao and Argos method  
FEATURE:  
NAME/KEY: potential Transmembrane helix  
LOCATION: 4..20  
IDENTIFICATION METHOD: Klein, Kanehisa and Delisi method  
FEATURE:  
NAME/KEY: potential Transmembrane helix  
LOCATION: 4..24  
IDENTIFICATION METHOD: Eisenberg, Schwarz, Komaromy  
IDENTIFICATION METHOD: and Wall method  
FEATURE:  
NAME/KEY: potential N-myristoylation site  
LOCATION: 12  
IDENTIFICATION METHOD: prosite match  
FEATURE:  
NAME/KEY: potential Transmembrane helix  
LOCATION: 50..70  
IDENTIFICATION METHOD: Eisenberg, Schwarz, Komaromy and Wall method  
FEATURE:  
NAME/KEY: potential N-glycosylation site  
LOCATION: 57

IDENTIFICATION METHOD: prosite match  
FEATURE:  
NAME/KEY: potential Transmembrane helix  
LOCATION: 76..96  
IDENTIFICATION METHOD: Eisenberg, Schwarz, Komaromy and Wall method  
FEATURE:  
NAME/KEY: potential Tyrosine kinase phosphorylation site  
LOCATION: 78  
IDENTIFICATION METHOD: prosite match  
FEATURE:  
NAME/KEY: potential Casein kinase II phosphorylation site  
LOCATION: 84  
IDENTIFICATION METHOD: prosite match  
FEATURE:  
NAME/KEY: potential Leucine zipper pattern  
LOCATION: 94..115  
IDENTIFICATION METHOD: prosite match  
FEATURE:  
NAME/KEY: potential N-myristoylation site  
LOCATION: 119  
IDENTIFICATION METHOD: prosite match  
FEATURE:  
NAME/KEY: potential Protein kinase C phosphorylation site  
LOCATION: 133  
IDENTIFICATION METHOD: prosite match  
FEATURE:  
NAME/KEY: potential Casein kinase II phosphorylation site  
LOCATION: 147  
IDENTIFICATION METHOD: prosite match  
FEATURE:  
NAME/KEY: potential Protein kinase C phosphorylation site  
LOCATION: 194  
IDENTIFICATION METHOD: prosite match  
FEATURE:  
NAME/KEY: potential Tyrosine kinase phosphorylation site  
LOCATION: 215  
IDENTIFICATION METHOD: prosite match  
FEATURE:  
NAME/KEY: potential Tyrosine sulfatation site  
LOCATION: 221  
IDENTIFICATION METHOD: prosite match  
FEATURE:  
NAME/KEY: potential cAMP- and cGMP-dependent protein kinase phosphorylation site  
LOCATION: 233  
IDENTIFICATION METHOD: prosite match  
FEATURE:  
NAME/KEY: potential Casein kinase II phosphorylation site  
LOCATION: 235  
IDENTIFICATION METHOD: prosite match  
FEATURE:  
NAME/KEY: potential Protein kinase C phosphorylation site  
LOCATION: 306  
IDENTIFICATION METHOD: prosite match  
FEATURE:  
NAME/KEY: potential Transmembrane helix  
LOCATION: 310..330  
IDENTIFICATION METHOD: Eisenberg, Schwarz, Komaromy and Wall method  
FEATURE:  
NAME/KEY: potential N-myristoylation site  
LOCATION: 319  
IDENTIFICATION METHOD: prosite match  
FEATURE:  
NAME/KEY: potential N-myristoylation site  
LOCATION: 323  
IDENTIFICATION METHOD: prosite match  
FEATURE:  
NAME/KEY: potential Amidation site  
LOCATION: 329  
IDENTIFICATION METHOD: prosite match  
FEATURE:  
NAME/KEY: potential Transmembrane helix  
LOCATION: 333..353  
IDENTIFICATION METHOD: Eisenberg, Schwarz, Komaromy and Wall method

FEATURE:  
NAME/KEY: potential N-myristoylation site  
LOCATION: 341  
IDENTIFICATION METHOD: prosite match  
FEATURE:  
NAME/KEY: potential protein kinase C phosphorylation site  
LOCATION: 350  
IDENTIFICATION METHOD: prosite match  
US-08-996-306-4

Query Match 97.3%; Score 1871; DB 2; Length 35;:  
Best Local Similarity 100.0%; Pred. No. 2, 2e-194;  
Matches 353; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 12 MRYLLPSVVLGTAFTYVLAMGVRLSAPFLPARFYQALDRLYCVYOSWLFEEFENTG 71  
DB 1 MRYLLPSVVLGTAFTYVLAMGVRLSAPFLPARFYQALDRLYCVYOSWLFEEFENTG 60  
QY 72 VOILLYDGLPRKNENIYYLANHOSVDWIVADIIAIRONALGHVRYVLEGLKMLPLXGC 131  
DB 61 VOILLYDGLPRKNENIYYLANHOSVDWIVADIIAIRONALGHVRYVLEGLKMLPLXGC 120  
QY 132 YFAOHGGIYVRSAPFKEMKRNKLSQSYVDAQTPMYLVIPEGRYRNPEDQTKVLSASQAF 191  
DB 121 YFAOHGGIYVRSAPFKEMKRNKLSQSYVDAQTPMYLVIPEGRYRNPEDQTKVLSASQAF 180  
QY 192 AAGRGVLVLRHVLTPTIKATVAFDCMKKNYLDALYDVTVVYEGKDDGQRRSPTEFL 251  
DB 181 AAGRGVLVLRHVLTPTIKATVAFDCMKKNYLDALYDVTVVYEGKDDGQRRSPTEFL 240  
QY 252 CKECPKIHIDRIDKDVPEEOHMRRLHERFEIKDKMLIEFESPDPERRRPFGKS 311  
DB 241 CKECPKIHIDRIDKDVPEEOHMRRLHERFEIKDKMLIEFESPDPERRRPFGKS 300  
QY 312 VNSKLSIKKTLPSMLIISGLTAGMLTDAGRKLYVNTWYIGTLGCLMWTIKA 364  
DB 301 VNSKLSIKKTLPSMLIISGLTAGMLTDAGRKLYVNTWYIGTLGCLMWTIKA 353

## RESULT 5

US-09-338-907-4  
Sequence 4, Application US/09338907  
Patent No. 6265546  
GENERAL INFORMATION:  
APPLICANT: Cohen, Daniel  
APPLICANT: Blumenfeld, Marta  
APPLICANT: Ilya, Chumakov  
APPLICANT: Bougueret, Lydie  
TITLE OF INVENTION: PROSTATE CANCER GENE  
FILE REFERENCE: GENSET.18CPICP  
CURRENT APPLICATION NUMBER: US/09/338.907  
CURRENT FILING DATE: 1999-06-23  
EARLIER APPLICATION NUMBER: 08/996.306  
EARLIER FILING DATE: 1997-12-22  
EARLIER APPLICATION NUMBER: 60/099.658  
EARLIER FILING DATE: 1998-09-09  
EARLIER APPLICATION NUMBER: 09/218.207  
EARLIER FILING DATE: 1998-12-22  
NUMBER OF SEQ ID NOS: 578  
SOFTWARE: Patent.pm  
SEQ ID NO 4  
LENGTH: 353  
TYPE: PRT  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: HELIX  
LOCATION: 1..33  
OTHER INFORMATION: Rao and Argos identification method, potential helix  
FEATURE:  
NAME/KEY: HELIX  
LOCATION: 4..20  
OTHER INFORMATION: Klein, Kanehisa and Delist identification method, potential

FEATURE:  
NAME/KEY: HELIX  
LOCATION: 4..24  
OTHER INFORMATION: Eisenberg, Schwarz, Komarony, Wall identification method,  
FEATURE:  
NAME/KEY: MYRISTATE  
LOCATION: 12..16  
OTHER INFORMATION: prosite match  
FEATURE:  
NAME/KEY: HELIX  
LOCATION: 50..70  
OTHER INFORMATION: Eisenberg, Schwarz, Komarony, Wall identification method,  
FEATURE:  
NAME/KEY: CARBOHYD  
LOCATION: 57..59  
OTHER INFORMATION: prosite match  
FEATURE:  
NAME/KEY: HELIX  
LOCATION: 76..96  
OTHER INFORMATION: Eisenberg, Schwarz, Komarony, Wall identification method,  
FEATURE:  
NAME/KEY: PHOSPHORYLATION  
LOCATION: 78  
OTHER INFORMATION: potential Tyrosine kinase site, prosite match  
FEATURE:  
NAME/KEY: PHOSPHORYLATION  
LOCATION: 84  
OTHER INFORMATION: potential caseine kinase II site, prosite match  
FEATURE:  
NAME/KEY: SITE  
LOCATION: 94..115  
OTHER INFORMATION: potential leucine zipper site, prosite match  
FEATURE:  
NAME/KEY: MYRISTATE  
LOCATION: 119..123  
OTHER INFORMATION: potential site, prosite match  
FEATURE:  
NAME/KEY: PHOSPHORYLATION  
LOCATION: 133  
OTHER INFORMATION: potential protein kinase C, prosite match  
FEATURE:  
NAME/KEY: PHOSPHORYLATION  
LOCATION: 147  
OTHER INFORMATION: potential caseine kinase II site, prosite match  
FEATURE:  
NAME/KEY: PHOSPHORYLATION  
LOCATION: 194  
OTHER INFORMATION: potential protein kinase C, prosite match  
FEATURE:  
NAME/KEY: PHOSPHORYLATION  
LOCATION: 215  
OTHER INFORMATION: potential Tyrosine kinase site, prosite match  
FEATURE:  
NAME/KEY: SULFATATION  
LOCATION: 221  
OTHER INFORMATION: prosite match  
FEATURE:  
NAME/KEY: PHOSPHORYLATION  
LOCATION: 233  
OTHER INFORMATION: potential CAMP and cGMP dependant protein kinase site,  
FEATURE:  
NAME/KEY: PHOSPHORYLATION  
LOCATION: 235  
OTHER INFORMATION: potential caseine kinase II site, prosite match  
FEATURE:  
NAME/KEY: PHOSPHORYLATION  
LOCATION: 306  
OTHER INFORMATION: potential protein kinase C, prosite match  
FEATURE:  
NAME/KEY: HELIX  
LOCATION: 310..330  
OTHER INFORMATION: Eisenberg, Schwarz, Komarony, Wall identification method,  
FEATURE:

```
/ NAME/KEY: MYRISTATE
/ LOCATION: 319..323
/ OTHER INFORMATION: Prosite match
/ FEATURE:
/ NAME/KEY: MYRISTATE
/ LOCATION: 323..327
/ OTHER INFORMATION: Prosite match
/ FEATURE:
/ NAME/KEY: AMIDATION
/ LOCATION: 329
/ OTHER INFORMATION: Prosite match
/ FEATURE:
/ NAME/KEY: HELIX
/ LOCATION: 333..353
/ OTHER INFORMATION: Eisenberg, Schwarz, Komarony, Wall identification method, F
/ FEATURE:
/ NAME/KEY: MYRISTATE
/ LOCATION: 341..345
/ OTHER INFORMATION: Prosite match
/ FEATURE:
/ NAME/KEY: PHOSPHORYLATION
/ LOCATION: 350
/ OTHER INFORMATION: potential protein kinase C, Prosite match
/ US-09-338-907-4
```

Query Match 97.2%; Score 1871; DB 4; Length 353;

Best Local Similarity 100.0%; Pred. No. 2.2e-194; Mismatches 0; Indels 0; Gaps 0;

Matches 353; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY 12 MRYLLPSVVLGTAPTYVYVANGVWRLLSAFLPARFYQALDDRLCYCYOSMVLFFPENYTG 71
DB 1 MRYLLPSVVLGTAPTYVYVANGVWRLLSAFLPARFYQALDDRLCYCYOSMVLFFPENYTG 60
QY 72 VOILLYGDLPKKENIITYLANHOSYVDWIVADILAIQNALGHVRYVLEKGLKWLPLYGC 131
DB 61 VOILLYGDLPKKENIITYLANHOSYVDWIVADILAIQNALGHVRYVLEKGLKWLPLYGC 120
QY 132 YFAOHGGIYVRSKAKFNEKEMRNKLOSVDAGTPMYTVLFFPGSTRNPEQTVLSAQAF 191
DB 121 YFAOHGGIYVRSKAKFNEKEMRNKLOSVDAGTPMYTVLFFPGSTRNPEQTVLSAQAF 180
QY 192 AAQRGLAVLKHYLPRIKATHYAFDCMKNYLDAIYDVTVYVYEGKDDGGORRESPTTEFL 251
DB 181 AAQRGLAVLKHYLPRIKATHYAFDCMKNYLDAIYDVTVYVYEGKDDGGORRESPTTEFL 240
QY 232 CKECPRIHIHIDRIKDDVPEOEHRMRLHERFEIKDKMLIEFYESPDPERRKRPFGKS 311
DB 241 CKECPRIHIHIDRIKDDVPEOEHRMRLHERFEIKDKMLIEFYESPDPERRKRPFGKS 300
QY 312 VNSKLSIKTKTLPMSMLISGLTAGMLMTDAGRKLVTWNTVYTGLLGCLMTTICA 364
DB 301 VNSKLSIKTKTLPMSMLISGLTAGMLMTDAGRKLVTWNTVYTGLLGCLMTTICA 353
```

#### RESULT 6

```
US-09-218-207-4
/ Sequence 4, Application US/09218207
/ Patent No. 6346381
/ GENERAL INFORMATION:
/ APPLICANT: Cohen, Daniel
/ APPLICANT: Blumenfeld, Marta
/ APPLICANT: Ilyia, Chumakov
/ APPLICANT: Bougueleret, Lydie
/ TITLE OF INVENTION: Prostate cancer gene
/ FILE REFERENCE: GENSET.018CPI
/ CURRENT APPLICATION NUMBER: US/09/218.207
/ CURRENT FILING DATE: 1998-12-22
/ EARLIER APPLICATION NUMBER: 08/996.306
/ EARLIER FILING DATE: 1997-12-22
/ EARLIER APPLICATION NUMBER: 60/099.658
/ EARLIER FILING DATE: 1998-09-09
/ NUMBER OF SEQ ID NOS: 578
```

```
/ SOFTWARE: Patent.pm
/ SEQ ID NO 4
/ LENGTH: 353
/ TYPE: PRT
/ ORGANISM: Homo sapiens
/ FEATURE:
/ NAME/KEY: HELIX
/ LOCATION: 1..33
/ OTHER INFORMATION: Rao and Argos identification method, potential helix
/ FEATURE:
/ NAME/KEY: HELIX
/ LOCATION: 4..20
/ OTHER INFORMATION: Klein, Kanehisa and Delisi identification method, potent
/ FEATURE:
/ NAME/KEY: HELIX
/ LOCATION: 4..24
/ OTHER INFORMATION: Eisenberg, Schwarz, Komarony, Wall identification method,
/ FEATURE:
/ NAME/KEY: MYRISTATE
/ LOCATION: 12..16
/ OTHER INFORMATION: Prosite match
/ FEATURE:
/ NAME/KEY: HELIX
/ LOCATION: 50..70
/ OTHER INFORMATION: Eisenberg, Schwarz, Komarony, Wall identification method,
/ FEATURE:
/ NAME/KEY: CARBOHYD
/ LOCATION: 57..59
/ OTHER INFORMATION: Prosite match
/ FEATURE:
/ NAME/KEY: HELIX
/ LOCATION: 76..96
/ OTHER INFORMATION: Eisenberg, Schwarz, Komarony, Wall identification method,
/ FEATURE:
/ NAME/KEY: PHOSPHORYLATION
/ LOCATION: 78
/ OTHER INFORMATION: potential Tyrosine kinase site, Prosite match
/ FEATURE:
/ NAME/KEY: PHOSPHORYLATION
/ LOCATION: 84
/ OTHER INFORMATION: potential caseine kinase II site, Prosite match
/ FEATURE:
/ NAME/KEY: SITE
/ LOCATION: 94..115
/ OTHER INFORMATION: potential Leucine zipper site, Prosite match
/ FEATURE:
/ NAME/KEY: MYRISTATE
/ LOCATION: 119..123
/ OTHER INFORMATION: potential site, Prosite match
/ FEATURE:
/ NAME/KEY: PHOSPHORYLATION
/ LOCATION: 133
/ OTHER INFORMATION: potential protein kinase C, Prosite match
/ FEATURE:
/ NAME/KEY: PHOSPHORYLATION
/ LOCATION: 147
/ OTHER INFORMATION: potential caseine kinase II site, Prosite match
/ FEATURE:
/ NAME/KEY: PHOSPHORYLATION
/ LOCATION: 194
/ OTHER INFORMATION: potential protein kinase C, Prosite match
/ FEATURE:
/ NAME/KEY: PHOSPHORYLATION
/ LOCATION: 215
/ OTHER INFORMATION: potential Tyrosine kinase site, Prosite match
/ FEATURE:
/ NAME/KEY: SULFATATION
/ LOCATION: 221
/ OTHER INFORMATION: Prosite match
/ FEATURE:
/ NAME/KEY: PHOSPHORYLATION
/ LOCATION: 233
/ OTHER INFORMATION: potential CAMP and cGMP dependant protein kinase site,
```





```

; EARLIER APPLICATION NUMBER: 08/996,306
; EARLIER FILING DATE: 1997-12-22
; EARLIER APPLICATION NUMBER: 60/099,658
; EARLIER FILING DATE: 1998-09-09
; NUMBER OF SEQ ID NOS: 578
; SOFTWARE: Patent.pm
; SEQ ID NO 134
; LENGTH: 315
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: 81..83
; OTHER INFORMATION: Box 1
; FEATURE:
; NAME/KEY: SITE
; LOCATION: 160..165
; OTHER INFORMATION: Box 11
US-09-218-207-134
```

```

Query Match          85.4%; Score 1644; DB 4; Length 315;
Best Local Similarity 89.2%; Pred. No. 7.4e-170;
Matches 315; Conservative 0; Mismatches 0; Indels 38; Gaps 1;
```

```

QY 12 MRYLLPSVLLGTAFTYVLAAGVWRLSAFLPARFYQALDDRLCYQSQAVLFEFFENYTG 71
DB 1 MRYLLPSVLLGTAFTYVLAAGVWRLSAFLPARFYQALDDRLCYQSQAVLFEFFENYTG 60
QY 72 VOILLVGDLPRKNENIYYLANHOSYDWMIVADILAIQNALGHRVYLKGLKMLPLYGC 131
DB 61 VOILLVGDLPRKNENIYYLANHOSYDWMIVADILAIQNALGHRVYLKGLKMLPLYGC 120
QY 132 YFAOHGSIYVKRSKAKFNEKEMRNKLSYVDAGTPMYLVIFPEGTRNPEQTKVLSASQAF 191
DB 121 YFAOHGSIYVKRSKAKFNEKEMRNKLSYVDAGTPMYLVIFPEGTRNPEQTKVLSASQAF 180
QY 192 AAORGSLAVLKHVLTPIKATHVAFDCMKNYLDIYDVTYVYEGKDDGGQRESPTMTFEL 251
DB 181 AAO-----GKDDGGQRESPTMTFEL 202
QY 252 CKECPEKIHIDRIDKKDVEEOEHMRMLHERFEIKDKMLIEFYESPDERRRKRPFGKS 311
DB 203 CKECPEKIHIDRIDKKDVEEOEHMRMLHERFEIKDKMLIEFYESPDERRRKRPFGKS 262
QY 312 VNSKLSIKRTLPMSMLISGLTAGMLMTDAGRKLYVTWITWYIGTLGCLMTVITKA 364
DB 263 VNSKLSIKRTLPMSMLISGLTAGMLMTDAGRKLYVTWITWYIGTLGCLMTVITKA 315
```

```

RESULT 9
US-09-338-907-74
; Sequence 74, Application US/09338907
; Patent No. 6265546
; GENERAL INFORMATION:
; APPLICANT: Cohen, Daniel
; APPLICANT: Blumenfeld, Marta
; APPLICANT: Ilya, Chumakov
; APPLICANT: Bougueleret, Lydie
; TITLE OF INVENTION: PROSTATE CANCER GENE
; FILE REFERENCE: GENSET.18CP1CP
; CURRENT APPLICATION NUMBER: US/09/338,907
; EARLIER APPLICATION NUMBER: 08/996,306
; EARLIER FILING DATE: 1997-12-22
; EARLIER APPLICATION NUMBER: 60/099,658
; EARLIER FILING DATE: 1998-09-09
; EARLIER APPLICATION NUMBER: 09/218,207
; EARLIER FILING DATE: 1998-12-22
; NUMBER OF SEQ ID NOS: 578
; SOFTWARE: Patent.pm
; SEQ ID NO 74
; LENGTH: 354
```

```

; TYPE: PRT
; ORGANISM: Mus musculus
US-09-338-907-74
```

```

Query Match          81.5%; Score 1568.5; DB 4; Length 354;
Best Local Similarity 81.4%; Pred. No. 1.4e-161;
Matches 288; Conservative 39; Mismatches 26; Indels 1; Gaps 1;
```

```

QY 12 MRYLLPSVLLGTAFTYVLAAGVWRLSAFLPARFYQALDDRLCYQSQAVLFEFFENYTG 71
DB 1 MRYLLPSVLLGTAFTYVLAAGVWRLSAFLPARFYQALDDRLCYQSQAVLFEFFENYTG 60
QY 72 VOILLVGDLPRKNENIYYLANHOSYDWMIVADILAIQNALGHRVYLKGLKMLPLYGC 131
DB 61 VOILLVGDLPRKNENIYYLANHOSYDWMIVADILAIQNALGHRVYLKGLKMLPLYGC 120
QY 132 YFAOHGSIYVKRSKAKFNEKEMRNKLSYVDAGTPMYLVIFPEGTRNPEQTKVLSASQAF 191
DB 121 YFAOHGSIYVKRSKAKFNEKEMRNKLSYVDAGTPMYLVIFPEGTRNPEQTKVLSASQAF 180
QY 192 AAORGSLAVLKHVLTPIKATHVAFDCMKNYLDIYDVTYVYEGKDDG-GQRESPTMTFEL 250
DB 181 AAORGSLAVLKHVLTPIKATHVAFDCMKNYLDIYDVTYVYEGKDDG-GQRESPTMTFEL 240
QY 251 CKECPEKIHIDRIDKKDVEEOEHMRMLHERFEIKDKMLIEFYESPDERRRKRPFGK 310
DB 241 CKECPEKIHIDRIDKKDVEEOEHMRMLHERFEIKDKMLIEFYESPDERRRKRPFGK 300
QY 311 VNSKLSIKRTLPMSMLISGLTAGMLMTDAGRKLYVTWITWYIGTLGCLMTVITKA 364
DB 301 VNSKLSIKRTLPMSMLISGLTAGMLMTDAGRKLYVTWITWYIGTLGCLMTVITKA 354
```

```

RESULT 10
US-09-218-207-74
; Sequence 74, Application US/09218207
; Patent No. 6346381
; GENERAL INFORMATION:
; APPLICANT: Cohen, Daniel
; APPLICANT: Blumenfeld, Marta
; APPLICANT: Ilya, Chumakov
; APPLICANT: Bougueleret, Lydie
; TITLE OF INVENTION: Prostate cancer gene
; FILE REFERENCE: GENSET.018CP1
; CURRENT APPLICATION NUMBER: US/09/218,207
; EARLIER APPLICATION NUMBER: 08/996,306
; EARLIER FILING DATE: 1997-12-22
; EARLIER APPLICATION NUMBER: 60/099,658
; EARLIER FILING DATE: 1998-09-09
; NUMBER OF SEQ ID NOS: 578
; SOFTWARE: Patent.pm
; SEQ ID NO 74
; LENGTH: 354
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-218-207-74
```

```

Query Match          81.5%; Score 1568.5; DB 4; Length 354;
Best Local Similarity 81.4%; Pred. No. 1.4e-161;
Matches 288; Conservative 39; Mismatches 26; Indels 1; Gaps 1;

QY 12 MRYLLPSVLLGTAFTYVLAAGVWRLSAFLPARFYQALDDRLCYQSQAVLFEFFENYTG 71
DB 1 MRYLLPSVLLGTAFTYVLAAGVWRLSAFLPARFYQALDDRLCYQSQAVLFEFFENYTG 60
QY 72 VOILLVGDLPRKNENIYYLANHOSYDWMIVADILAIQNALGHRVYLKGLKMLPLYGC 131
DB 61 VOILLVGDLPRKNENIYYLANHOSYDWMIVADILAIQNALGHRVYLKGLKMLPLYGC 120
QY 132 YFAOHGSIYVKRSKAKFNEKEMRNKLSYVDAGTPMYLVIFPEGTRNPEQTKVLSASQAF 191
```

```
|||||
Db 121 YFAOHGSIYVRSKAFENDKEMRSKLOSVDAGTPTMYLVIFPEEGTRYNATYTKILSASQAF 180
Qy 192 AAORGGLAVLHVLPRIKATHVAFDCKKNYLDAIYDVTYVEGDDG-GQRESPTMEF 250
|||
Db 181 AAORGGLAVLHVLPRIKATHVAFDCKSHLDATYDVTYVEGEGKSGKSNPSMTEF 240
Qy 251 LCKECPRIHIHIDRIKDVPEDEQEHMRWLHEREFIKDKMLIEFESPPDERRRKRFPGK 310
|||
Db 241 LCKGCPRLHIHFDRIIDNEVPEDEQEHMKMLHEREFIKDKMLIEFYDPSDERRRKRFPGK 300
Qy 311 SVNSKLSIKTKLPMSMLISGLTAGMLMTDAGRKLVTNTWITGTLGCLMTVITKA 364
|||
Db 301 SVHSRLSVKTKLPMSVLISGLTAGMLMTESGRKLYMGTMVITGTLGCLMTVITKA 354

RESULT 11
US-09-338-907-135
; Sequence 135, Application US/09338907
; Patent No. 6265546
; GENERAL INFORMATION:
; APPLICANT: Cohen, Daniel
; APPLICANT: Blumenfeld, Marla
; APPLICANT: Ilya, Chumakov
; APPLICANT: Bougueleret, Lydie
; TITLE OF INVENTION: PROSTATE CANCER GENE
; FILE REFERENCE: GENSET.18CP1CP
; CURRENT APPLICATION NUMBER: US/09/338,907
; EARLIER FILING DATE: 1999-06-23
; EARLIER APPLICATION NUMBER: 08/996,306
; EARLIER FILING DATE: 1997-12-22
; EARLIER APPLICATION NUMBER: 60/099,658
; EARLIER FILING DATE: 1998-09-09
; EARLIER APPLICATION NUMBER: 09/218,207
; EARLIER FILING DATE: 1998-12-22
; NUMBER OF SEQ ID NOS: 578
; SOFTWARE: Patent.pm
; SEQ ID NO 135
; LENGTH: 300
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: 81..83
; OTHER INFORMATION: Box I
; FEATURE:
; NAME/KEY: SITE
; LOCATION: 160..165
; OTHER INFORMATION: Box II
US-09-338-907-135

Query Match 80.8%; Score 1555.5; DB 4; Length 300;
Best Local Similarity 85.0%; Pred. No.2.7e-160;
Matches 300; Conservative 0; Mismatches 0; Indels 53; Gaps 1;
```

```
|||||
Db 188 CKECPRIHIHIDRIKDVPEDEQEHMRWLHEREFIKDKMLIEFYDPSDERRRKRFPGKS 247
Qy 312 VNSKLSIKTKLPMSMLISGLTAGMLMTDAGRKLVTNTWITGTLGCLMTVITKA 364
|||
Db 248 VNSKLSIKTKLPMSMLISGLTAGMLMTDAGRKLVTNTWITGTLGCLMTVITKA 300

RESULT 12
US-09-218-207-135
; Sequence 135, Application US/09218207
; Patent No. 6346381
; GENERAL INFORMATION:
; APPLICANT: Cohen, Daniel
; APPLICANT: Blumenfeld, Marla
; APPLICANT: Ilya, Chumakov
; APPLICANT: Bougueleret, Lydie
; TITLE OF INVENTION: Prostate cancer gene
; FILE REFERENCE: GENSET.018CP1
; CURRENT APPLICATION NUMBER: US/09/218,207
; EARLIER FILING DATE: 1998-12-22
; EARLIER APPLICATION NUMBER: 08/996,306
; EARLIER FILING DATE: 1997-12-22
; EARLIER APPLICATION NUMBER: 60/099,658
; EARLIER FILING DATE: 1998-09-09
; NUMBER OF SEQ ID NOS: 578
; SOFTWARE: Patent.pm
; SEQ ID NO 135
; LENGTH: 300
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: 81..83
; OTHER INFORMATION: Box I
; FEATURE:
; NAME/KEY: SITE
; LOCATION: 160..165
; OTHER INFORMATION: Box II
US-09-218-207-135

Query Match 80.8%; Score 1555.5; DB 4; Length 300;
Best Local Similarity 85.0%; Pred. No.2.7e-160;
Matches 300; Conservative 0; Mismatches 0; Indels 53; Gaps 1;
```

```
; Sequence 127, Application US/09338907
; Patent No. 6265546
; GENERAL INFORMATION:
; APPLICANT: Cohen, Daniel
; APPLICANT: Blumenfeld, Marta
; APPLICANT: Ilya, Chumakov
; APPLICANT: Bougueleret, Lydie
; TITLE OF INVENTION: PROSTATE CANCER GENE
; FILE REFERENCE: GENSET.18CPLCP
; CURRENT APPLICATION NUMBER: US/09/338,907
; CURRENT FILING DATE: 1999-06-23
; EARLIER APPLICATION NUMBER: 08/996,306
; EARLIER FILING DATE: 1997-12-22
; EARLIER APPLICATION NUMBER: 60/099,658
; EARLIER FILING DATE: 1998-09-09
; EARLIER APPLICATION NUMBER: 09/218,207
; EARLIER FILING DATE: 1998-12-22
; NUMBER OF SEQ ID NOS: 578
; SOFTWARE: Patent.pm
; SEQ ID NO 127
; LENGTH: 291
; TYPE: PRF
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: 98..103
; OTHER INFORMATION: Box II
; FEATURE:
; NAME/KEY: SITE
; LOCATION: 149..157
; OTHER INFORMATION: Box III
US-09-338-907-127
```

```
Query Match      77.9%; Score 1498; DB 4; Length 291;
Best Local Similarity 82.4%; Pred. No. 4,4e-154;
Matches 291; Conservative 0; Mismatches 0; Indels 62; Gaps 1;

QY 12 MRLLPSVLLGTAAPYVLAAGVWRLLSAFLPARFYQALDRLCYVOSWLPFFENYTG 71
DB 1 MRLLPSVLLGTAAPYVLAAGVWRLLSAFLPARFYQALDRLCYVOSWLPFFENYTG 60
QY 72 VOILLGDLPRKNENIYYLANHSTVDIVADIIAIRONALGHVRYLKEGLKMLPLYGC 131
DB 61 V----- 61
QY 132 YFAOHGIIYVRSKAKFENEKEMRNKLOSVDAGTPTMYLVIPEEGTRYNPEOTKVLASQAF 191
DB 62 ---QHGGIYVRSKAKFENEKEMRNKLOSVDAGTPTMYLVIPEEGTRYNPEOTKVLASQAF 118
QY 192 AAGRGALVLAHVLPRIKATHVAFDCMKNYLDAIYDVTVVYEGKDDGGQRESPTMTTEFL 251
DB 119 AAGRGALVLAHVLPRIKATHVAFDCMKNYLDAIYDVTVVYEGKDDGGQRESPTMTTEFL 178
QY 252 CKECPKIHIDRIDKDDVPEOEHHMRWLHERFEIKDKMLIEFYESPDPERRRRPFQGS 311
DB 179 CKECPKIHIDRIDKDDVPEOEHHMRWLHERFEIKDKMLIEFYESPDPERRRRPFQGS 238
QY 312 VNSKLSIKKTLPSMLILSGLTAGMLMTDAGRKLYVNTWYIGTLGCLMTVITKA 364
DB 239 VNSKLSIKKTLPSMLILSGLTAGMLMTDAGRKLYVNTWYIGTLGCLMTVITKA 291
```

```
RESULT 14
US-09-218-207-127
; Sequence 127, Application US/09218207
; Patent No. 6346381
; GENERAL INFORMATION:
; APPLICANT: Cohen, Daniel
; APPLICANT: Blumenfeld, Marta
; APPLICANT: Ilya, Chumakov
; APPLICANT: Bougueleret, Lydie
; TITLE OF INVENTION: Prostate cancer gene
```

```
; FILE REFERENCE: GENSET.018CPI
; CURRENT APPLICATION NUMBER: US/09/218,207
; CURRENT FILING DATE: 1998-12-22
; EARLIER APPLICATION NUMBER: 08/996,306
; EARLIER FILING DATE: 1997-12-22
; EARLIER APPLICATION NUMBER: 60/099,658
; EARLIER FILING DATE: 1998-09-09
; NUMBER OF SEQ ID NOS: 578
; SOFTWARE: Patent.pm
; SEQ ID NO 127
; LENGTH: 291
; TYPE: PRF
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: 98..103
; OTHER INFORMATION: Box II
; FEATURE:
; NAME/KEY: SITE
; LOCATION: 149..157
; OTHER INFORMATION: Box III
US-09-218-207-127
```

```
Query Match      77.9%; Score 1498; DB 4; Length 291;
Best Local Similarity 82.4%; Pred. No. 4,4e-154;
Matches 291; Conservative 0; Mismatches 0; Indels 62; Gaps 1;

QY 12 MRLLPSVLLGTAAPYVLAAGVWRLLSAFLPARFYQALDRLCYVOSWLPFFENYTG 71
DB 1 MRLLPSVLLGTAAPYVLAAGVWRLLSAFLPARFYQALDRLCYVOSWLPFFENYTG 60
QY 72 VOILLGDLPRKNENIYYLANHSTVDIVADIIAIRONALGHVRYLKEGLKMLPLYGC 131
DB 61 V----- 61
QY 132 YFAOHGIIYVRSKAKFENEKEMRNKLOSVDAGTPTMYLVIPEEGTRYNPEOTKVLASQAF 191
DB 62 ---QHGGIYVRSKAKFENEKEMRNKLOSVDAGTPTMYLVIPEEGTRYNPEOTKVLASQAF 118
QY 192 AAGRGALVLAHVLPRIKATHVAFDCMKNYLDAIYDVTVVYEGKDDGGQRESPTMTTEFL 251
DB 119 AAGRGALVLAHVLPRIKATHVAFDCMKNYLDAIYDVTVVYEGKDDGGQRESPTMTTEFL 178
QY 252 CKECPKIHIDRIDKDDVPEOEHHMRWLHERFEIKDKMLIEFYESPDPERRRRPFQGS 311
DB 179 CKECPKIHIDRIDKDDVPEOEHHMRWLHERFEIKDKMLIEFYESPDPERRRRPFQGS 238
QY 312 VNSKLSIKKTLPSMLILSGLTAGMLMTDAGRKLYVNTWYIGTLGCLMTVITKA 364
DB 239 VNSKLSIKKTLPSMLILSGLTAGMLMTDAGRKLYVNTWYIGTLGCLMTVITKA 291
```

```
RESULT 15
US-09-338-907-128
; Sequence 128, Application US/09338907
; Patent No. 6265546
; GENERAL INFORMATION:
; APPLICANT: Cohen, Daniel
; APPLICANT: Blumenfeld, Marta
; APPLICANT: Ilya, Chumakov
; APPLICANT: Bougueleret, Lydie
; TITLE OF INVENTION: PROSTATE CANCER GENE
; FILE REFERENCE: GENSET.18CPLCP
; CURRENT APPLICATION NUMBER: US/09/338,907
; CURRENT FILING DATE: 1999-06-23
; EARLIER APPLICATION NUMBER: 08/996,306
; EARLIER FILING DATE: 1997-12-22
; EARLIER APPLICATION NUMBER: 60/099,658
; EARLIER FILING DATE: 1998-09-09
; EARLIER APPLICATION NUMBER: 09/218,207
; EARLIER FILING DATE: 1998-12-22
; NUMBER OF SEQ ID NOS: 578
```

```
; SOFTWARE: Patent.pm
; SEQ ID NO 128
; LENGTH: 261
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: 68..73
; OTHER INFORMATION: Box II
; FEATURE:
; NAME/KEY: SITE
; LOCATION: 119..127
; OTHER INFORMATION: Box III
; US-09-338-907-128
```

```
Query Match          68.9%; Score 1325; DB 4; length 261:
Best Local Similarity 73.9%; Pred. No. 2,1e-135;
Matches 261; Conservative 0; Mismatches 0; Indels 92; Gaps 1;
```

```
QY 12 MRYLLPSVYLLGTAPTYVLANGVRLLSAFLPARFYQALDDRLCYQSMVLPFFENTYG 71
    ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1 MRYLLPSVYLLGTAPTYVLANGVRLLSAFLPARFYQALDDRLCYQSMVLPFFENTYG 60

QY 72 VQILLYGDLPRNKENIITYLANHGSTVDWIVADILAIRQNALGHVRYVLKEGIKMLPLYGC 131
    ||
Db 61 VQ----- 62

QY 132 YFAQHGGIYVRSKAFNKEKEMRNKLQSYVDAGTWPVLYIPEEGTRYNPQOTKVLASQAF 191
    ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 63 -----MYLVIPEEGTRYNPQOTKVLASQAF 88

QY 192 AAQGLAVLKHLVTPRIKATHVAFDCMKNYLDAYDVYVYEGKDDGGQRRSPPTMEFL 251
    ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 89 AAQGLAVLKHLVTPRIKATHVAFDCMKNYLDAYDVYVYEGKDDGGQRRSPPTMEFL 148

QY 252 CKECPKIHIDRIDDKDVPPEQEHMRMLHEREFIKDKMLIEFYESPDPERRKRFPQGS 311
    ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 149 CKECPKIHIDRIDDKDVPPEQEHMRMLHEREFIKDKMLIEFYESPDPERRKRFPQGS 208

QY 312 VNSKLSIKKTLPSMLISGLTFAGMLMTDAGRKLVTNTWYIGTLGCLMWTIKA 364
    ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 209 VNSKLSIKKTLPSMLISGLTFAGMLMTDAGRKLVTNTWYIGTLGCLMWTIKA 261
```

Search completed: August 28, 2002, 11:08:11  
Job time: 290 sec

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compen Ltd.

OM protein - protein search, using sw model

Run on: August 28, 2002, 11:14:10 ; Search time 306.96 Seconds  
(without alignments)  
417.386 Million cell updates/sec

Title: US-09-853-526-5  
Sequence: 1 MLSTLVHTYSMRYLPSVY.....YVNTWIVGTLCGLWTFIKA 364

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 3502263 seqs, 351980561 residues  
Total number of hits satisfying chosen parameters: 3502263

Minimum DB seq length: 0  
Maximum DB seq length: 200000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Pending\_Patents\_AA\_Main:\*  
1: /cgn2\_6/ptodata/2/paa/PCrus\_COMB.pep:\*  
2: /cgn2\_6/ptodata/2/paa/US06\_COMB.pep:\*  
3: /cgn2\_6/ptodata/2/paa/US07\_COMB.pep:\*  
4: /cgn2\_6/ptodata/2/paa/US08\_COMB.pep:\*  
5: /cgn2\_6/ptodata/2/paa/US081\_COMB.pep:\*  
6: /cgn2\_6/ptodata/2/paa/US082\_COMB.pep:\*  
7: /cgn2\_6/ptodata/2/paa/US083\_COMB.pep:\*  
8: /cgn2\_6/ptodata/2/paa/US084\_COMB.pep:\*  
9: /cgn2\_6/ptodata/2/paa/US085\_COMB.pep:\*  
10: /cgn2\_6/ptodata/2/paa/US086\_COMB.pep:\*  
11: /cgn2\_6/ptodata/2/paa/US087\_COMB.pep:\*  
12: /cgn2\_6/ptodata/2/paa/US088\_COMB.pep:\*  
13: /cgn2\_6/ptodata/2/paa/US089\_COMB.pep:\*  
14: /cgn2\_6/ptodata/2/paa/US090\_COMB.pep:\*  
15: /cgn2\_6/ptodata/2/paa/US091\_COMB.pep:\*  
16: /cgn2\_6/ptodata/2/paa/US092\_COMB.pep:\*  
17: /cgn2\_6/ptodata/2/paa/US093\_COMB.pep:\*  
18: /cgn2\_6/ptodata/2/paa/US094\_COMB.pep:\*  
19: /cgn2\_6/ptodata/2/paa/US095\_COMB.pep:\*  
20: /cgn2\_6/ptodata/2/paa/US096\_COMB.pep:\*  
21: /cgn2\_6/ptodata/2/paa/US097\_COMB.pep:\*  
22: /cgn2\_6/ptodata/2/paa/US098\_COMB.pep:\*  
23: /cgn2\_6/ptodata/2/paa/US099\_COMB.pep:\*  
24: /cgn2\_6/ptodata/2/paa/US100\_COMB.pep:\*  
25: /cgn2\_6/ptodata/2/paa/US101\_COMB.pep:\*  
26: /cgn2\_6/ptodata/2/paa/US60\_COMB.pep:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
1	1924	100.0	364	18	US-09-436-919-1
2	1924	100.0	364	22	US-09-817-910-7
3	1924	100.0	364	22	US-09-853-526-5
4	1924	100.0	364	23	US-09-901-484A-5
5	1924	100.0	364	26	US-60-099-658-5
6	1913	99.4	372	18	US-09-488-725A-6308
7	1871	97.2	353	1	PCT-US01-01431-59

8	1871	97.2	353	1	PCT-US01-01431-79	Sequence 79, Appl
9	1871	97.2	353	1	PCT-US01-11988-1474	Sequence 1474, Ap
10	1871	97.2	353	22	PCT-US01-11988-1475	Sequence 1475, Ap
11	1871	97.2	353	22	US-09-833-245-1474	Sequence 1474, Ap
12	1871	97.2	353	22	US-09-833-245-1475	Sequence 1475, Ap
13	1871	97.2	353	22	US-09-853-526-4	Sequence 4, Appl
14	1871	97.2	353	23	US-09-901-484A-4	Sequence 4, Appl
15	1871	97.2	353	23	US-09-915-582-59	Sequence 59, Appl
16	1871	97.2	353	23	US-09-915-582-79	Sequence 79, Appl
17	1871	97.2	353	26	US-60-099-658-4	Sequence 4, Appl
18	1861	96.7	353	18	US-09-488-725A-2736	Sequence 2736, Ap
19	1644	85.4	315	22	US-09-853-526-134	Sequence 134, App
20	1644	85.4	315	23	US-09-901-484A-134	Sequence 110, App
21	1619.5	84.2	450	26	US-60-245-222-110	Sequence 74, Appl
22	1568.5	81.5	354	22	US-09-853-526-74	Sequence 74, Appl
23	1568.5	81.5	354	23	US-09-901-484A-74	Sequence 74, Appl
24	1568.5	81.5	354	26	US-60-099-658-74	Sequence 309, Appl
25	1559.5	81.1	1032	26	US-60-212-413-309	Sequence 248, App
26	1559.5	81.1	1032	26	US-60-225-518-248	Sequence 135, App
27	1555.5	80.8	300	22	US-09-853-526-135	Sequence 135, App
28	1555.5	80.8	300	23	US-09-901-484A-135	Sequence 127, App
29	1498	77.9	291	22	US-09-853-526-127	Sequence 127, App
30	1498	77.9	291	23	US-09-901-484A-127	Sequence 47, Appl
31	1427	74.2	269	1	PCT-US01-01327-47	Sequence 128, App
32	1325	68.9	261	22	US-09-853-526-128	Sequence 70, Appl
33	1325	68.9	261	23	US-09-901-484A-128	Sequence 70, Appl
34	1203	62.5	228	22	US-09-853-526-70	Sequence 70, Appl
35	1203	62.5	228	23	US-09-901-484A-70	Sequence 126, App
36	1203	62.5	228	26	US-60-099-658-70	Sequence 126, App
37	1182.5	61.5	238	22	US-09-853-526-126	Sequence 136, App
38	1182.5	61.5	238	23	US-09-901-484A-126	Sequence 133, App
39	976	50.7	185	22	US-09-853-526-136	Sequence 133, App
40	976	50.7	185	23	US-09-901-484A-136	Sequence 132, App
41	938	48.8	176	18	US-09-436-919-5	Sequence 5, Appl
42	820	42.6	182	22	US-09-853-526-133	Sequence 133, App
43	820	42.6	182	23	US-09-901-484A-133	Sequence 132, App
44	449	23.3	97	22	US-09-853-526-132	Sequence 132, App
45	449	23.3	97	23	US-09-901-484A-132	Sequence 132, App

ALIGNMENTS

RESULT 1  
US-09-436-919-1  
; Sequence 1, Application US/09436919A  
; GENERAL INFORMATION:  
; APPLICANT: Leung, David W  
; TITLE OF INVENTION: Human Lysophosphatidic Acid Acyltransferase-epsilon  
; FILE REFERENCE: 1801B  
; CURRENT APPLICATION NUMBER: US/09/436,919A  
; CURRENT FILING DATE: 1999-11-09  
; NUMBER OF SEQ ID NOS: 5  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 1  
; LENGTH: 364  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; FEATURE:  
; OTHER INFORMATION: hLPAepsilon  
US-09-436-919-1

Query Match 100.0%; Score 1924; DB 18; Length 364;  
Best local Similarity 100.0%; Pred. No. 4e-189;  
Matches 364; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 MLSTLVHTYSMRYLPSVLTGTAFTYVLAMGVMLSLPAPRYQALDRLCYVOS 60  
DB 1 MLSTLVHTYSMRYLPSVLTGTAFTYVLAMGVMLSLPAPRYQALDRLCYVOS 60  
QY 61 MVLFFENTGVQVILLYGDLPKKRNIIYLANQSTVDIVADILATRONALGHVRYVLK 120  
|||||

```
Db 61 MVLFFENYTGVOILLYGDLPKKNENIITYLANHOSTVDWIVADILAIROMALGHVRYLK 120
Qy 121 EGLKMLPLGYCYFAOHGCIYVKSRAKFNEMKRNKLOSVDAGTAPMYLVIFPESTRYNPE 180
Db 121 EGLKMLPLGYCYFAOHGCIYVKSRAKFNEMKRNKLOSVDAGTAPMYLVIFPESTRYNPE 180
Qy 181 QTKVLSASQAFAAORGLAVLKHVLTLPRIKATHVAFDCMKNYLDAIYDVTVYVECKDDGGQ 240
Db 181 QTKVLSASQAFAAORGLAVLKHVLTLPRIKATHVAFDCMKNYLDAIYDVTVYVECKDDGGQ 240
Qy 241 RRESPTMEFLCKECPRIHIHIDRIDKKDVEEEOEHMRMLHEFEIKDKMLIEFYESPD 300
Db 241 RRESPTMEFLCKECPRIHIHIDRIDKKDVEEEOEHMRMLHEFEIKDKMLIEFYESPD 300
Qy 301 PERKRRPGKSVNSKLSIKKTLPSMLILSGITAGMLMTDAGRKLKYVMTWITIGTLGCLMW 360
Db 301 PERKRRPGKSVNSKLSIKKTLPSMLILSGITAGMLMTDAGRKLKYVMTWITIGTLGCLMW 360
Qy 361 TIRKA 364
Db 361 TIRKA 364
```

```
RESULT 2
US-09-817-910-7
; Sequence 7, Application US/09817910
; GENERAL INFORMATION:
; APPLICANT: Meyers, Rachel A.
; APPLICANT: Macbeth, Kyle J.
; APPLICANT: Williamson, Mark
; APPLICANT: Rudolph-Owen, Laura A.
; APPLICANT: Tsai, Fong-Ying
; TITLE OF INVENTION: 46743 and 27417 NOVEL HUMAN
; FILE REFERENCE: 10448-032001
; CURRENT APPLICATION NUMBER: US/09/817, 910
; CURRENT FILING DATE: 2001-03-26
; PRIOR APPLICATION NUMBER: US 60/192,092
; PRIOR FILING DATE: 2000-03-24
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 7
; LENGTH: 364
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-817-910-7
```

```
Query Match 100.0%; Score 1924; DB 22; Length 364;
Best Local Similarity 100.0%; Pred. No. 4e-189;
Matches 364; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Qy 1 MLSTLVHTYSMRLLPSVYLLGTAPTYVLAWGWRLLSAFLPARFYQALDDRLCYQYS 60
Db 1 MLSTLVHTYSMRLLPSVYLLGTAPTYVLAWGWRLLSAFLPARFYQALDDRLCYQYS 60
Qy 61 MVLFFENYTGVOILLYGDLPKKNENIITYLANHOSTVDWIVADILAIROMALGHVRYLK 120
Db 61 MVLFFENYTGVOILLYGDLPKKNENIITYLANHOSTVDWIVADILAIROMALGHVRYLK 120
Qy 121 EGLKMLPLGYCYFAOHGCIYVKSRAKFNEMKRNKLOSVDAGTAPMYLVIFPESTRYNPE 180
Db 121 EGLKMLPLGYCYFAOHGCIYVKSRAKFNEMKRNKLOSVDAGTAPMYLVIFPESTRYNPE 180
Qy 181 QTKVLSASQAFAAORGLAVLKHVLTLPRIKATHVAFDCMKNYLDAIYDVTVYVECKDDGGQ 240
Db 181 QTKVLSASQAFAAORGLAVLKHVLTLPRIKATHVAFDCMKNYLDAIYDVTVYVECKDDGGQ 240
Qy 241 RRESPTMEFLCKECPRIHIHIDRIDKKDVEEEOEHMRMLHEFEIKDKMLIEFYESPD 300
Db 241 RRESPTMEFLCKECPRIHIHIDRIDKKDVEEEOEHMRMLHEFEIKDKMLIEFYESPD 300
Qy 301 PERKRRPGKSVNSKLSIKKTLPSMLILSGITAGMLMTDAGRKLKYVMTWITIGTLGCLMW 360
Db 301 PERKRRPGKSVNSKLSIKKTLPSMLILSGITAGMLMTDAGRKLKYVMTWITIGTLGCLMW 360
```

```
Db 301 PERKRRPGKSVNSKLSIKKTLPSMLILSGITAGMLMTDAGRKLKYVMTWITIGTLGCLMW 360
Qy 361 TIRKA 364
Db 361 TIRKA 364
```

```
RESULT 3
US-09-853-526-5
; Sequence 5, Application US/09853526
; GENERAL INFORMATION:
; APPLICANT: Cohen, Daniel
; APPLICANT: Blumenfeld, Marta
; APPLICANT: Ilya, Chumakov
; APPLICANT: Bougueleret, Lydie
; TITLE OF INVENTION: PROSTATE CANCER GENE
; FILE REFERENCE: GENSET.18CPICP
; CURRENT APPLICATION NUMBER: US/09/853, 526
; CURRENT FILING DATE: 2001-05-11
; PRIOR APPLICATION NUMBER: 09/338,907
; PRIOR FILING DATE: 1999-06-23
; PRIOR APPLICATION NUMBER: 08/996,306
; PRIOR FILING DATE: 1997-12-22
; PRIOR APPLICATION NUMBER: 60/099,658
; PRIOR FILING DATE: 1998-09-09
; PRIOR APPLICATION NUMBER: 09/218,207
; PRIOR FILING DATE: 1998-12-22
; NUMBER OF SEQ ID NOS: 578
; SOFTWARE: Patent.pm
; SEQ ID NO 5
; LENGTH: 364
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-853-526-5
```

```
Query Match 100.0%; Score 1924; DB 22; Length 364;
Best Local Similarity 100.0%; Pred. No. 4e-189;
Matches 364; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Qy 1 MLSTLVHTYSMRLLPSVYLLGTAPTYVLAWGWRLLSAFLPARFYQALDDRLCYQYS 60
Db 1 MLSTLVHTYSMRLLPSVYLLGTAPTYVLAWGWRLLSAFLPARFYQALDDRLCYQYS 60
Qy 61 MVLFFENYTGVOILLYGDLPKKNENIITYLANHOSTVDWIVADILAIROMALGHVRYLK 120
Db 61 MVLFFENYTGVOILLYGDLPKKNENIITYLANHOSTVDWIVADILAIROMALGHVRYLK 120
Qy 121 EGLKMLPLGYCYFAOHGCIYVKSRAKFNEMKRNKLOSVDAGTAPMYLVIFPESTRYNPE 180
Db 121 EGLKMLPLGYCYFAOHGCIYVKSRAKFNEMKRNKLOSVDAGTAPMYLVIFPESTRYNPE 180
Qy 181 QTKVLSASQAFAAORGLAVLKHVLTLPRIKATHVAFDCMKNYLDAIYDVTVYVECKDDGGQ 240
Db 181 QTKVLSASQAFAAORGLAVLKHVLTLPRIKATHVAFDCMKNYLDAIYDVTVYVECKDDGGQ 240
Qy 241 RRESPTMEFLCKECPRIHIHIDRIDKKDVEEEOEHMRMLHEFEIKDKMLIEFYESPD 300
Db 241 RRESPTMEFLCKECPRIHIHIDRIDKKDVEEEOEHMRMLHEFEIKDKMLIEFYESPD 300
Qy 301 PERKRRPGKSVNSKLSIKKTLPSMLILSGITAGMLMTDAGRKLKYVMTWITIGTLGCLMW 360
Db 301 PERKRRPGKSVNSKLSIKKTLPSMLILSGITAGMLMTDAGRKLKYVMTWITIGTLGCLMW 360
Qy 361 TIRKA 364
Db 361 TIRKA 364
```

```
RESULT 4
US-09-901-484A-5
; Sequence 5, Application US/09901484A
```

```

; GENERAL INFORMATION:
; APPLICANT: Cohen, Daniel
; APPLICANT: Blumenfeld, Marta
; APPLICANT: Chumakov, Ilya
; APPLICANT: Bougueleret, Lydie
; TITLE OF INVENTION: Prostate Cancer Gene
; FILE REFERENCE: GEN-T11XC3D2
; CURRENT APPLICATION NUMBER: US/09/901,484A
; CURRENT FILING DATE: 2001-07-09
; PRIOR APPLICATION NUMBER: US 08/996,306
; PRIOR FILING DATE: 1997-12-22
; PRIOR APPLICATION NUMBER: US 60/099,658
; PRIOR FILING DATE: 1998-09-09
; PRIOR APPLICATION NUMBER: US 09/218,207
; PRIOR FILING DATE: 1998-12-22
; PRIOR APPLICATION NUMBER: US 09/338,907
; PRIOR FILING DATE: 1999-06-23
; PRIOR APPLICATION NUMBER: US 09/853,526
; PRIOR FILING DATE: 2001-05-11
; NUMBER OF SEQ ID NOS: 578
; SOFTWARE: Patentln version 3.1
; SEQ ID NO: 5
; LENGTH: 364
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-901-484A-5

```

```

Query Match      100.0%; Score 1924; DB 23; Length 364;
Best Local Similarity 100.0%; Pred. No. 4e-189;
Matches 364; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1 MLISLVHTYSMRYLLPSVLLGTAPTYVLAMGVMRLLSAFLPARFYQALDRLCYVOS 60
DB 1 MLISLVHTYSMRYLLPSVLLGTAPTYVLAMGVMRLLSAFLPARFYQALDRLCYVOS 60
QY 61 MVLFFENTYGVQILLYGDLPRKKNITLYLANHSTVDMIVADILAIRONALGHVRYVK 120
DB 61 MVLFFENTYGVQILLYGDLPRKKNITLYLANHSTVDMIVADILAIRONALGHVRYVK 120
QY 121 EGIKMLPLVGCYFAOHGIIYVRSKAFNEKEMNKLOSVDAGTPMYLVIFFEGTRYNPE 180
DB 121 EGIKMLPLVGCYFAOHGIIYVRSKAFNEKEMNKLOSVDAGTPMYLVIFFEGTRYNPE 180
QY 181 OFKVLASQAFAAQRGLAVLKHVLPRIKATHVAFDCKMNYDAIYDVTVYVEGKDDGQ 240
DB 181 OFKVLASQAFAAQRGLAVLKHVLPRIKATHVAFDCKMNYDAIYDVTVYVEGKDDGQ 240
QY 241 RRESPTMTEFLCKECKPKIHIDRIDKDVPEOEHRMRMLHERFEIKDKMLIEFYESP 300
DB 241 RRESPTMTEFLCKECKPKIHIDRIDKDVPEOEHRMRMLHERFEIKDKMLIEFYESP 300
QY 301 PERRKRPFGKSVNSKLSIKKTLPMSMLISGLTAGMLMTDAGRKLVTYNTWYIGTLGCLMV 360
DB 301 PERRKRPFGKSVNSKLSIKKTLPMSMLISGLTAGMLMTDAGRKLVTYNTWYIGTLGCLMV 360
QY 361 TTKA 364
DB 361 TTKA 364

```

```

RESULT 5
US-60-099-658-5
; Sequence 5, Application US/60099658
; GENERAL INFORMATION:
; APPLICANT: Cohen, Daniel
; APPLICANT: Chumakov, Ilya
; APPLICANT: Blumenfeld, Marta
; APPLICANT: Bougueleret, Lydie
; TITLE OF INVENTION: Prostate cancer gene
; NUMBER OF SEQUENCES: 99
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Knobbe, Martens, Olson & Bear

```

```

; STREET: 501 West Broadway
; CITY: San Diego
; STATE: California
; COUNTRY: USA
; ZIP: 92101-3505
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: Win95
; SOFTWARE: Word
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/60/099,658
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Israelsen, Ned A.
; REGISTRATION NUMBER: 29,655
; REFERENCE/DOCKET NUMBER: GENSET.018APR
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (619) 235-8550
; TELEFAX: (619) 235-0176
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 364 amino acids
; TYPE: AMINO ACID
; STRANDEDNESS: SINGLE
; TOPOLOGY: LINEAR
; MOLECULE TYPE: PROTEIN
; ORIGINAL SOURCE:
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: potential protein
; LOCATION: 1..364
US-60-099-658-5

```

```

Query Match      100.0%; Score 1924; DB 26; Length 364;
Best Local Similarity 100.0%; Pred. No. 4e-189;
Matches 364; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1 MLISLVHTYSMRYLLPSVLLGTAPTYVLAMGVMRLLSAFLPARFYQALDRLCYVOS 60
DB 1 MLISLVHTYSMRYLLPSVLLGTAPTYVLAMGVMRLLSAFLPARFYQALDRLCYVOS 60
QY 61 MVLFFENTYGVQILLYGDLPRKKNITLYLANHSTVDMIVADILAIRONALGHVRYVK 120
DB 61 MVLFFENTYGVQILLYGDLPRKKNITLYLANHSTVDMIVADILAIRONALGHVRYVK 120
QY 121 EGIKMLPLVGCYFAOHGIIYVRSKAFNEKEMNKLOSVDAGTPMYLVIFFEGTRYNPE 180
DB 121 EGIKMLPLVGCYFAOHGIIYVRSKAFNEKEMNKLOSVDAGTPMYLVIFFEGTRYNPE 180
QY 181 OFKVLASQAFAAQRGLAVLKHVLPRIKATHVAFDCKMNYDAIYDVTVYVEGKDDGQ 240
DB 181 OFKVLASQAFAAQRGLAVLKHVLPRIKATHVAFDCKMNYDAIYDVTVYVEGKDDGQ 240
QY 241 RRESPTMTEFLCKECKPKIHIDRIDKDVPEOEHRMRMLHERFEIKDKMLIEFYESP 300
DB 241 RRESPTMTEFLCKECKPKIHIDRIDKDVPEOEHRMRMLHERFEIKDKMLIEFYESP 300
QY 301 PERRKRPFGKSVNSKLSIKKTLPMSMLISGLTAGMLMTDAGRKLVTYNTWYIGTLGCLMV 360
DB 301 PERRKRPFGKSVNSKLSIKKTLPMSMLISGLTAGMLMTDAGRKLVTYNTWYIGTLGCLMV 360
QY 361 TTKA 364
DB 361 TTKA 364

```

```

RESULT 6
US-09-488-725A-6308
; Sequence 6308, Application US/09488725A
; GENERAL INFORMATION:

```

```
; APPLICANT: Hyseq, Inc
; TITLE OF INVENTION: Novel Nucleic Acid and Polypeptides
; FILE REFERENCE: 784FLPCT
; CURRENT APPLICATION NUMBER: US/09/488, 725A
; CURRENT FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: US/09/488, 725
; PRIOR FILING DATE: 2000-01-21
; PRIOR APPLICATION NUMBER: US09/552, 317
; PRIOR FILING DATE: 2000-04-25
; PRIOR APPLICATION NUMBER: US09/598, 042
; PRIOR FILING DATE: 2000-06-20
; PRIOR APPLICATION NUMBER: US09/620, 312
; PRIOR FILING DATE: 2000-07-19
; PRIOR APPLICATION NUMBER: US09/653, 450
; PRIOR FILING DATE: 2000-08-31
; PRIOR APPLICATION NUMBER: US09/662, 191
; PRIOR FILING DATE: 2000-09-14
; PRIOR APPLICATION NUMBER: US09/693, 036
; PRIOR FILING DATE: 2000-10-19
; PRIOR APPLICATION NUMBER: US09/727, 344
; PRIOR FILING DATE: 2000-11-29
; NUMBER OF SEQ ID NOS: 7144
; SOFTWARE: pt_FL_genes_b Versions 1.0
; SEQ ID NO 6308
; LENGTH: 372
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-488-725A-6308

Query Match      99.4%; Score 1913; DB 18; Length 372;
Best Local Similarity 99.7%; Pred. No. 5.7e-188;
Matches 363; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 MLSLVLTHTYSMRVLLPSVVLGTAPTYVLAMGWRLLSAFLPARFYQALDDRLCYQSVYOS 60
DB 9 MLSLVLTHTYSMRVLLPSVVLGTAPTYVLAMGWRLLSAFLPARFYQALDDRLCYQSVYOS 68
OY 61 MVLFFENYTGVOILLYGLDLPKKNENITYLANHOSYVDWIVADILAIKQNALGHVRVYLK 120
DB 69 MVLFFENYTGVOILLYGLDLPKKNENITYLANHOSYVDWIVADILAIKQNALGHVRVYLK 128
OY 121 EGLWMLPLVLCYFPAOHGIIYVRSKAKENEMKRNKLOSVDAGTPMYLVIFPEGTRYNPE 180
DB 129 EGLWMLPLVLCYFPAOHGIIYVRSKAKENEMKRNKLOSVDAGTPMYLVIFPEGTRYNPE 188
OY 181 QTKVLSASQAFPAOAGRLAVLKHVLTLPRIKATHVAFDCMKNYLDIYDVTYVYEGKDDGGO 240
DB 189 QTKVLSASQAFPAOAGRLAVLKHVLTLPRIKATHVAFDCMKNYLDIYDVTYVYEGKDDGGO 248
OY 241 RRESPMTEFLCKECPKIHIDRIDDKDVPDEQEHMRMWLHERPEIKDKMLIEFYES3PD 300
DB 249 RRESPMTEFLCKECPKIHIDRIDDKDVPDEQEHMRMWLHERPEIKDKMLIEFYES3PD 308
OY 301 PERKKRPPGKSVNSKLSIKTKLPSMLLSGLTAGMLMTDAGRKLIVNTWIIYGTLLGCLMW 360
DB 309 PERKKRPPGKSVNSKLSIKTKLPSMLLSGLTAGMLMTDAGRKLIVNTWIIYGTLLGCLMW 368
OY 361 TIRKA 364
DB 369 TIRKA 372

RESULT 7
PCT-US01-01431-59
; Sequence 59, Application PC/TUS0101431
; GENERAL INFORMATION:
; APPLICANT: Human Genome Sciences, Inc.
; TITLE OF INVENTION: 17 human secreted proteins
; FILE REFERENCE: PS723PCT
; CURRENT APPLICATION NUMBER: PCT/US01/01431
; CURRENT FILING DATE: 2001-01-17
; PRIOR APPLICATION NUMBER: 60/179, 065
```

```
; PRIOR FILING DATE: 2000-01-31
; PRIOR APPLICATION NUMBER: 60/180, 628
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: 60/231, 968
; PRIOR FILING DATE: 2000-09-12
; NUMBER OF SEQ ID NOS: 97
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 59
; LENGTH: 353
; TYPE: PRT
; ORGANISM: Homo sapiens
PCT-US01-01431-59

Query Match      97.2%; Score 1871; DB 1; Length 353;
Best Local Similarity 100.0%; Pred. No. 1.1e-183;
Matches 353; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 12 MRYLLPSVVLGTAPTYVLAMGWRLLSAFLPARFYQALDDRLCYQSVYOSMVLFFENYTG 71
DB 1 MRYLLPSVVLGTAPTYVLAMGWRLLSAFLPARFYQALDDRLCYQSVYOSMVLFFENYTG 60
OY 72 VOILLYGLDLPKKNENITYLANHOSYVDWIVADILAIKQNALGHVRVYLKKGKMLPLVYGC 121
DB 61 VOILLYGLDLPKKNENITYLANHOSYVDWIVADILAIKQNALGHVRVYLKKGKMLPLVYGC 120
OY 132 YFAOHGIIYVRSKAKENEMKRNKLOSVDAGTPMYLVIFPEGTRYNPEOTKVLASQAF 191
DB 121 YFAOHGIIYVRSKAKENEMKRNKLOSVDAGTPMYLVIFPEGTRYNPEOTKVLASQAF 180
OY 192 AAORGGLAVLKHVLTLPRIKATHVAFDCMKNYLDIYDVTYVYEGKDDGGO RRESPMTEFL 251
DB 181 AAORGGLAVLKHVLTLPRIKATHVAFDCMKNYLDIYDVTYVYEGKDDGGO RRESPMTEFL 240
OY 252 CKECSPKIHIDRIDDKDVPDEQEHMRMWLHERPEIKDKMLIEFYESPDERRRRRPPGKS 311
DB 241 CKECSPKIHIDRIDDKDVPDEQEHMRMWLHERPEIKDKMLIEFYESPDERRRRRPPGKS 300
OY 312 VNSKLSIKTKLPSMLLSGLTAGMLMTDAGRKLIVNTWIIYGTLLGCLMW TIRKA 364
DB 301 VNSKLSIKTKLPSMLLSGLTAGMLMTDAGRKLIVNTWIIYGTLLGCLMW TIRKA 353

RESULT 8
PCT-US01-01431-79
; Sequence 79, Application PC/TUS0101431
; GENERAL INFORMATION:
; APPLICANT: Human Genome Sciences, Inc.
; TITLE OF INVENTION: 17 human secreted proteins
; FILE REFERENCE: PS723PCT
; CURRENT APPLICATION NUMBER: PCT/US01/01431
; CURRENT FILING DATE: 2001-01-17
; PRIOR APPLICATION NUMBER: 60/179, 065
; PRIOR FILING DATE: 2000-01-31
; PRIOR APPLICATION NUMBER: 60/180, 628
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: 60/231, 968
; PRIOR FILING DATE: 2000-09-12
; NUMBER OF SEQ ID NOS: 97
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 79
; LENGTH: 353
; TYPE: PRT
; ORGANISM: Homo sapiens
PCT-US01-01431-79

Query Match      97.2%; Score 1871; DB 1; Length 353;
Best Local Similarity 100.0%; Pred. No. 1.1e-183;
Matches 353; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 12 MRYLLPSVVLGTAPTYVLAMGWRLLSAFLPARFYQALDDRLCYQSVYOSMVLFFENYTG 71
DB 1 MRYLLPSVVLGTAPTYVLAMGWRLLSAFLPARFYQALDDRLCYQSVYOSMVLFFENYTG 71
```



```
Db      1  MRLLPSVVLGTAPTYVLAAGWVRLLSAFLPARFYQALDDRLCYCYQSQWVLFEEFNYYTG 60
QY      72  VQILLYGDLPKKNENIYYLANHOSYDWMIVADILAIROMALGHVRYLKEGLKWLPLYGC 131
Db      61  VQILLYGDLPKKNENIYYLANHOSYDWMIVADILAIROMALGHVRYLKEGLKWLPLYGC 120
QY      132 YFAOHGGIYKRSKAKFNEKEMRNKLOSVDAGTPMYLVIFPESTRNPEQTKYLSASQAF 191
Db      121 YFAOHGGIYKRSKAKFNEKEMRNKLOSVDAGTPMYLVIFPESTRNPEQTKYLSASQAF 180
QY      192 AAORGLAVLKHVLPRIKATHVAFDCMKNYLDIYDVTVVYEGKDDGGQRESPTMTTEFL 251
Db      181 AAORGLAVLKHVLPRIKATHVAFDCMKNYLDIYDVTVVYEGKDDGGQRESPTMTTEFL 240
QY      252 CKCEPKIHIDRIDKKDVEEEOEHMRWMLHERFEIKDKMLIEFYSPDERRRKRPFGKS 311
Db      241 CKCEPKIHIDRIDKKDVEEEOEHMRWMLHERFEIKDKMLIEFYSPDERRRKRPFGKS 300
QY      312 VNSKLSIKTLPMSMLISGLTAGMLMTDAGRKLVTWNTWYIGTLGCLMWTIKA 364
Db      301 VNSKLSIKTLPMSMLISGLTAGMLMTDAGRKLVTWNTWYIGTLGCLMWTIKA 353
```

```
RESULT 9
PCT-US01-11988-1474
; Sequence 1474, Application PC/TUS0111988
```

```
; GENERAL INFORMATION:
```

```
; APPLICANT: Human Genome Sciences, Inc.
```

```
; TITLE OF INVENTION: Albumin Fusion Proteins
```

```
; FILE REFERENCE: PF546PCT
```

```
; CURRENT APPLICATION NUMBER: PCT/US01/11988
```

```
; CURRENT FILING DATE: 2001-01-12
```

```
; PRIOR APPLICATION NUMBER: 60/229, 358
```

```
; PRIOR FILING DATE: 2000-04-12
```

```
; PRIOR APPLICATION NUMBER: 60/256, 931
```

```
; PRIOR FILING DATE: 2000-12-21
```

```
; PRIOR APPLICATION NUMBER: 60/199, 384
```

```
; PRIOR FILING DATE: 2000-04-25
```

```
; NUMBER OF SEQ ID NOS: 2267
```

```
; SOFTWARE: PatentIn Ver. 2.1
```

```
; SEQ ID NO 1474
```

```
; LENGTH: 353
```

```
; TYPE: PRT
```

```
; ORGANISM: Homo sapiens
```

```
; PCT-US01-11988-1474
```

```
Query Match          97.2%; Score 1871; DB 1; Length 353;
Best Local Similarity 100.0%; Pred. No. 1.1e-183;
Matches 353; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      12  MRLLPSVVLGTAPTYVLAAGWVRLLSAFLPARFYQALDDRLCYCYQSQWVLFEEFNYYTG 71
Db      1  MRLLPSVVLGTAPTYVLAAGWVRLLSAFLPARFYQALDDRLCYCYQSQWVLFEEFNYYTG 60
QY      72  VQILLYGDLPKKNENIYYLANHOSYDWMIVADILAIROMALGHVRYLKEGLKWLPLYGC 131
Db      61  VQILLYGDLPKKNENIYYLANHOSYDWMIVADILAIROMALGHVRYLKEGLKWLPLYGC 120
QY      132 YFAOHGGIYKRSKAKFNEKEMRNKLOSVDAGTPMYLVIFPESTRNPEQTKYLSASQAF 191
Db      121 YFAOHGGIYKRSKAKFNEKEMRNKLOSVDAGTPMYLVIFPESTRNPEQTKYLSASQAF 180
QY      192 AAORGLAVLKHVLPRIKATHVAFDCMKNYLDIYDVTVVYEGKDDGGQRESPTMTTEFL 251
Db      181 AAORGLAVLKHVLPRIKATHVAFDCMKNYLDIYDVTVVYEGKDDGGQRESPTMTTEFL 240
QY      252 CKCEPKIHIDRIDKKDVEEEOEHMRWMLHERFEIKDKMLIEFYSPDERRRKRPFGKS 311
Db      241 CKCEPKIHIDRIDKKDVEEEOEHMRWMLHERFEIKDKMLIEFYSPDERRRKRPFGKS 300
QY      312 VNSKLSIKTLPMSMLISGLTAGMLMTDAGRKLVTWNTWYIGTLGCLMWTIKA 364
Db      301 VNSKLSIKTLPMSMLISGLTAGMLMTDAGRKLVTWNTWYIGTLGCLMWTIKA 353
```

```
Db      301 VNSKLSIKTLPMSMLISGLTAGMLMTDAGRKLVTWNTWYIGTLGCLMWTIKA 353
```

```
RESULT 10
PCT-US01-11988-1475
; Sequence 1475, Application PC/TUS0111988
```

```
; GENERAL INFORMATION:
```

```
; APPLICANT: Human Genome Sciences, Inc.
```

```
; TITLE OF INVENTION: Albumin Fusion Proteins
```

```
; FILE REFERENCE: PF546PCT
```

```
; CURRENT APPLICATION NUMBER: PCT/US01/11988
```

```
; CURRENT FILING DATE: 2001-01-12
```

```
; PRIOR APPLICATION NUMBER: 60/229, 358
```

```
; PRIOR FILING DATE: 2000-04-12
```

```
; PRIOR APPLICATION NUMBER: 60/256, 931
```

```
; PRIOR FILING DATE: 2000-12-21
```

```
; PRIOR APPLICATION NUMBER: 60/199, 384
```

```
; PRIOR FILING DATE: 2000-04-25
```

```
; NUMBER OF SEQ ID NOS: 2267
```

```
; SOFTWARE: PatentIn Ver. 2.1
```

```
; SEQ ID NO 1475
```

```
; LENGTH: 353
```

```
; TYPE: PRT
```

```
; ORGANISM: Homo sapiens
```

```
; PCT-US01-11988-1475
```

```
Query Match          97.2%; Score 1871; DB 1; Length 353;
Best Local Similarity 100.0%; Pred. No. 1.1e-183;
Matches 353; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      12  MRLLPSVVLGTAPTYVLAAGWVRLLSAFLPARFYQALDDRLCYCYQSQWVLFEEFNYYTG 71
Db      1  MRLLPSVVLGTAPTYVLAAGWVRLLSAFLPARFYQALDDRLCYCYQSQWVLFEEFNYYTG 60
QY      72  VQILLYGDLPKKNENIYYLANHOSYDWMIVADILAIROMALGHVRYLKEGLKWLPLYGC 131
Db      61  VQILLYGDLPKKNENIYYLANHOSYDWMIVADILAIROMALGHVRYLKEGLKWLPLYGC 120
QY      132 YFAOHGGIYKRSKAKFNEKEMRNKLOSVDAGTPMYLVIFPESTRNPEQTKYLSASQAF 191
Db      121 YFAOHGGIYKRSKAKFNEKEMRNKLOSVDAGTPMYLVIFPESTRNPEQTKYLSASQAF 180
QY      192 AAORGLAVLKHVLPRIKATHVAFDCMKNYLDIYDVTVVYEGKDDGGQRESPTMTTEFL 251
Db      181 AAORGLAVLKHVLPRIKATHVAFDCMKNYLDIYDVTVVYEGKDDGGQRESPTMTTEFL 240
QY      252 CKCEPKIHIDRIDKKDVEEEOEHMRWMLHERFEIKDKMLIEFYSPDERRRKRPFGKS 311
Db      241 CKCEPKIHIDRIDKKDVEEEOEHMRWMLHERFEIKDKMLIEFYSPDERRRKRPFGKS 300
QY      312 VNSKLSIKTLPMSMLISGLTAGMLMTDAGRKLVTWNTWYIGTLGCLMWTIKA 364
Db      301 VNSKLSIKTLPMSMLISGLTAGMLMTDAGRKLVTWNTWYIGTLGCLMWTIKA 353

RESULT 11
US-09-833-245-1474
; Sequence 1474, Application US/09833245
; GENERAL INFORMATION:
; APPLICANT: Human Genome Sciences, Inc.
; TITLE OF INVENTION: Albumin Fusion Proteins
; FILE REFERENCE: PF546PCT
; CURRENT APPLICATION NUMBER: US/09/833,245
; CURRENT FILING DATE: 2001-04-12
; PRIOR APPLICATION NUMBER: 60/229, 358
; PRIOR FILING DATE: 2000-04-12
; PRIOR APPLICATION NUMBER: 60/256, 931
; PRIOR FILING DATE: 2000-12-21
; PRIOR APPLICATION NUMBER: 60/199, 384
; PRIOR FILING DATE: 2000-04-25
; NUMBER OF SEQ ID NOS: 2267
; SOFTWARE: PatentIn Ver. 2.1
```

SEQ ID NO 1474  
LENGTH: 353  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-09-833-245-1474

Query Match 57.2%; Score 1871; DB 22; Length 353;  
Best Local Similarity 100.0%; Pred. No. 1,1e-183;  
Matches 353; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 12 MRYLLPSVLLGTAFTYVLAWGVRLSAPLPARYQALDRLCYVOSMVLFFFEENTG 71  
DB 1 MRYLLPSVLLGTAFTYVLAWGVRLSAPLPARYQALDRLCYVOSMVLFFFEENTG 60  
OY 72 VOILLYGLDLPKKNENIYLANHOSIVDWIVADILAIRONALGHVRYVLEKGLKMLPLXGC 131  
DB 61 VOILLYGLDLPKKNENIYLANHOSIVDWIVADILAIRONALGHVRYVLEKGLKMLPLXGC 120  
OY 132 YFAOHGGIYVRSKAFENKEKRNKLQSYVDAGTPMYLVIPEEGTRYNPEQTKVLSASQAF 191  
DB 121 YFAOHGGIYVRSKAFENKEKRNKLQSYVDAGTPMYLVIPEEGTRYNPEQTKVLSASQAF 180  
OY 192 AAOGLAVLKHLVLPRIKATHVAFDCKMKNYLDATYDVTVVYEGKDDGQORESPPTMEFL 251  
DB 181 AAOGLAVLKHLVLPRIKATHVAFDCKMKNYLDATYDVTVVYEGKDDGQORESPPTMEFL 240  
OY 252 CKECPKIHIDRIKDKVDPEQEHMRRWLHERFEIKDKMLIEFYESPDPERRRRFPKGS 311  
DB 241 CKECPKIHIDRIKDKVDPEQEHMRRWLHERFEIKDKMLIEFYESPDPERRRRFPKGS 300  
OY 312 VNSKLSIKTKLPSMLISGLTAGMLMTDAGRKLYVNTWYIGTLLGCLMWTIKA 364  
DB 301 VNSKLSIKTKLPSMLISGLTAGMLMTDAGRKLYVNTWYIGTLLGCLMWTIKA 353

RESULT 12  
US-09-833-245-1475  
Sequence 1475, Application US/09833245  
GENERAL INFORMATION:  
APPLICANT: Human Genome Sciences, Inc.  
TITLE OF INVENTION: Albumin Fusion Proteins  
FILE REFERENCE: P546PCT  
CURRENT APPLICATION NUMBER: US/09/833,245  
CURRENT FILING DATE: 2001-04-12  
PRIOR APPLICATION NUMBER: 60/229,358  
PRIOR FILING DATE: 2000-04-12  
PRIOR APPLICATION NUMBER: 60/256,931  
PRIOR FILING DATE: 2000-12-21  
PRIOR APPLICATION NUMBER: 60/199,384  
PRIOR FILING DATE: 2000-04-25  
NUMBER OF SEQ ID NOS: 2267  
SOFTWARE: Patentln Ver. 2.1  
SEQ ID NO 1475  
LENGTH: 353  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-09-833-245-1475

Query Match 97.2%; Score 1871; DB 22; Length 353;  
Best Local Similarity 100.0%; Pred. No. 1,1e-183;  
Matches 353; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 12 MRYLLPSVLLGTAFTYVLAWGVRLSAPLPARYQALDRLCYVOSMVLFFFEENTG 71  
DB 1 MRYLLPSVLLGTAFTYVLAWGVRLSAPLPARYQALDRLCYVOSMVLFFFEENTG 60  
OY 72 VOILLYGLDLPKKNENIYLANHOSIVDWIVADILAIRONALGHVRYVLEKGLKMLPLXGC 131  
DB 61 VOILLYGLDLPKKNENIYLANHOSIVDWIVADILAIRONALGHVRYVLEKGLKMLPLXGC 120  
OY 132 YFAOHGGIYVRSKAFENKEKRNKLQSYVDAGTPMYLVIPEEGTRYNPEQTKVLSASQAF 191

DB 121 YFAOHGGIYVRSKAFENKEKRNKLQSYVDAGTPMYLVIPEEGTRYNPEQTKVLSASQAF 180  
OY 192 AAOGLAVLKHLVLPRIKATHVAFDCKMKNYLDATYDVTVVYEGKDDGQORESPPTMEFL 251  
DB 181 AAOGLAVLKHLVLPRIKATHVAFDCKMKNYLDATYDVTVVYEGKDDGQORESPPTMEFL 240  
OY 252 CKECPKIHIDRIKDKVDPEQEHMRRWLHERFEIKDKMLIEFYESPDPERRRRFPKGS 311  
DB 241 CKECPKIHIDRIKDKVDPEQEHMRRWLHERFEIKDKMLIEFYESPDPERRRRFPKGS 300  
OY 312 VNSKLSIKTKLPSMLISGLTAGMLMTDAGRKLYVNTWYIGTLLGCLMWTIKA 364  
DB 301 VNSKLSIKTKLPSMLISGLTAGMLMTDAGRKLYVNTWYIGTLLGCLMWTIKA 353

RESULT 13  
US-09-853-526-4  
Sequence 4, Application US/09853526  
GENERAL INFORMATION:  
APPLICANT: Cohen, Daniel  
APPLICANT: Blumenfeld, Marta  
APPLICANT: Ilyu, Chumakov  
APPLICANT: Bougueleret, Lydie  
TITLE OF INVENTION: PROSTATE CANCER GENE  
FILE REFERENCE: GENSET.18CPICP  
CURRENT APPLICATION NUMBER: US/09/853,526  
CURRENT FILING DATE: 2001-05-11  
PRIOR APPLICATION NUMBER: 09/338,907  
PRIOR FILING DATE: 1999-06-23  
PRIOR APPLICATION NUMBER: 08/996,306  
PRIOR FILING DATE: 1997-12-22  
PRIOR APPLICATION NUMBER: 60/099,658  
PRIOR FILING DATE: 1998-09-09  
PRIOR APPLICATION NUMBER: 09/218,207  
PRIOR FILING DATE: 1998-12-22  
NUMBER OF SEQ ID NOS: 578  
SOFTWARE: Patent..pm  
SEQ ID NO 4  
LENGTH: 353  
TYPE: PRT  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: HELIX  
LOCATION: 1..33  
OTHER INFORMATION: Rao and Argos identification method, potential helix  
NAME/KEY: HELIX  
LOCATION: 4..20  
OTHER INFORMATION: Klein, Kanehisa and Delisi identification method, potent  
NAME/KEY: HELIX  
LOCATION: 4..24  
OTHER INFORMATION: Eisenberg, Schwarz, Komarony, Wall identification method,  
OTHER INFORMATION: helix  
NAME/KEY: MYRISTATE  
LOCATION: 12..16  
OTHER INFORMATION: Prosite match  
NAME/KEY: HELIX  
LOCATION: 50..70  
OTHER INFORMATION: Eisenberg, Schwarz, Komarony, Wall identification method,  
OTHER INFORMATION: helix  
NAME/KEY: CARBOHYD  
LOCATION: 57..59  
OTHER INFORMATION: Prosite match  
NAME/KEY: HELIX  
LOCATION: 76..96  
OTHER INFORMATION: Eisenberg, Schwarz, Komarony, Wall identification method,  
NAME/KEY: PHOSPHORYLATION  
LOCATION: 78  
OTHER INFORMATION: potential Tyrosine kinase site, Prosite match  
NAME/KEY: PHOSPHORYLATION  
LOCATION: 84  
OTHER INFORMATION: potential caseine kinase II site, Prosite match

```
NAME/KEY: SITE
LOCATION: 94..115
OTHER INFORMATION: potential Leucine zipper site, Prosite match
NAME/KEY: MYRISTATE
LOCATION: 119..123
OTHER INFORMATION: potential site, Prosite match
NAME/KEY: PHOSPHORYLATION
LOCATION: 133
OTHER INFORMATION: potential protein kinase C, Prosite match
NAME/KEY: PHOSPHORYLATION
LOCATION: 147
OTHER INFORMATION: potential caseine kinase II site, Prosite match
NAME/KEY: PHOSPHORYLATION
LOCATION: 194
OTHER INFORMATION: potential protein kinase C, Prosite match
NAME/KEY: PHOSPHORYLATION
LOCATION: 215
OTHER INFORMATION: potential Tyrosine kinase site, Prosite match
NAME/KEY: SULFATATION
LOCATION: 221
OTHER INFORMATION: Prosite match
NAME/KEY: PHOSPHORYLATION
LOCATION: 223
OTHER INFORMATION: potential cAMP and cGMP dependant protein kinase site, Prosite match
NAME/KEY: PHOSPHORYLATION
LOCATION: 235
OTHER INFORMATION: potential caseine kinase II site, Prosite match
NAME/KEY: PHOSPHORYLATION
LOCATION: 306
OTHER INFORMATION: potential protein kinase C, Prosite match
NAME/KEY: HELIX
LOCATION: 310..330
OTHER INFORMATION: Eisenberg, Schwarz, Komarony, Wall identification method, F
NAME/KEY: MYRISTATE
LOCATION: 319..323
OTHER INFORMATION: helix
NAME/KEY: MYRISTATE
LOCATION: 323..327
OTHER INFORMATION: Prosite match
NAME/KEY: MYRISTATE
LOCATION: 329
OTHER INFORMATION: Prosite match
NAME/KEY: AMIDATION
LOCATION: 329
OTHER INFORMATION: Prosite match
NAME/KEY: HELIX
LOCATION: 333..353
OTHER INFORMATION: Eisenberg, Schwarz, Komarony, Wall identification method, F
NAME/KEY: MYRISTATE
LOCATION: 341..345
OTHER INFORMATION: Prosite match
NAME/KEY: PHOSPHORYLATION
LOCATION: 350
OTHER INFORMATION: potential protein kinase C, Prosite match
US-09-853-526-4
```

```
Query Match 97.2%; Score 1871; DB 22; Length 353;
Best Local Similarity 100.0%; Pred. No. 1,1e-183;
Matches 353; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 12 MRLPLSPVLLGAPTYVLAAGVWRLLSAFLPARFYQALDDRLXYCYQSVLFEFFNYNG 71
DB 1 MRLPLSPVLLGAPTYVLAAGVWRLLSAFLPARFYQALDDRLXYCYQSVLFEFFNYNG 60
QY 72 VOLLTGDPKPKNKENTIIYLANHOSYVDWIVADILAIKRONALGHRVYLKGLKMLPLXGC 131
DB 61 VOLLTGDPKPKNKENTIIYLANHOSYVDWIVADILAIKRONALGHRVYLKGLKMLPLXGC 120
QY 132 YFAOHGIGIVKRSKAKNEKEMRNKLOSIVDAGTDMVLVIFPESTRYNEQOTKVLASQAF 191
DB 121 YFAOHGIGIVKRSKAKNEKEMRNKLOSIVDAGTDMVLVIFPESTRYNEQOTKVLASQAF 180
```

```
QY 192 AAGRGVLAVLKHVLTPIKATHVAFDCMKNYLAIYDVTVVYEGKDDGQRRSPMTTEFL 251
DB 181 AAGRGVLAVLKHVLTPIKATHVAFDCMKNYLAIYDVTVVYEGKDDGQRRSPMTTEFL 240
QY 252 CKCEPKIHIDRIDKKVDPEOEHRMRWLHEFEIKDKMLIEFYSPDERKRRPGKS 311
DB 241 CKCEPKIHIDRIDKKVDPEOEHRMRWLHEFEIKDKMLIEFYSPDERKRRPGKS 300
QY 312 VNSKLSIKTKPLPSMLILSGLTAGMLMTDAGRKLYVTWTIYGTLLGLMTYTKA 364
DB 301 VNSKLSIKTKPLPSMLILSGLTAGMLMTDAGRKLYVTWTIYGTLLGLMTYTKA 353
```

RESULT 14  
US-09-901-484A-4

```
Sequence 4, Application US/09901484A
GENERAL INFORMATION:
APPLICANT: Cohen, Daniel
APPLICANT: Blumenfeld, Marta
APPLICANT: Chumakov, Ilya
APPLICANT: Bougueleret, Lydie
TITLE OF INVENTION: Prostate Cancer Gene
FILE REFERENCE: GEN-T11XC3D2
CURRENT APPLICATION NUMBER: US/09/901,484A
PRIOR FILING DATE: 2001-07-09
PRIOR APPLICATION NUMBER: US 08/996,306
PRIOR FILING DATE: 1997-12-22
PRIOR APPLICATION NUMBER: US 60/099,658
PRIOR FILING DATE: 1998-09-09
PRIOR APPLICATION NUMBER: US 09/218,207
PRIOR FILING DATE: 1998-12-22
PRIOR APPLICATION NUMBER: US 09/338,907
PRIOR FILING DATE: 1999-06-23
PRIOR APPLICATION NUMBER: US 09/853,526
PRIOR FILING DATE: 2001-05-11
NUMBER OF SEQ ID NOS: 578
SOFTWARE: PatentIn version 3.1
SEQ ID NO 4
LENGTH: 353
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: HELIX
LOCATION: (1)..(33)
OTHER INFORMATION: Rao and Argos identification method, potential helix
NAME/KEY: HELIX
LOCATION: (4)..(20)
OTHER INFORMATION: Klein, Kanehisa and Delisi identification method, potential he
NAME/KEY: HELIX
LOCATION: (4)..(24)
OTHER INFORMATION: Eisenberg, Schwarz, Komarony, Wall identification method,
OTHER INFORMATION: potential helix
NAME/KEY: LIPID
LOCATION: (12)..(16)
OTHER INFORMATION: MYRISTATE, Prosite match
NAME/KEY: HELIX
LOCATION: (50)..(70)
OTHER INFORMATION: Eisenberg, Schwarz, Komarony, Wall identification method,
OTHER INFORMATION: potential helix
NAME/KEY: CARBOHYD
LOCATION: (57)..(59)
OTHER INFORMATION: Prosite match
NAME/KEY: HELIX
LOCATION: (76)..(96)
OTHER INFORMATION: Eisenberg, Schwarz, Komarony, Wall identification method,
OTHER INFORMATION: potential helix
NAME/KEY: MOD RES
LOCATION: (78)..(78)
OTHER INFORMATION: PHOSPHORYLATION, potential Tyrosine kinase site, Prosite match
NAME/KEY: MOD RES
LOCATION: (84)..(84)
OTHER INFORMATION: PHOSPHORYLATION, potential caseine kinase II site, Prosite mat
NAME/KEY: SITE
```

```
LOCATION: (94)..(115)
OTHER INFORMATION: Potential leucine zipper site, Prosite match
NAME/KEY: LIPID
LOCATION: (119)..(123)
OTHER INFORMATION: MYR::STATE, Prosite match
NAME/KEY: MOD_RES
LOCATION: (133)..(133)
OTHER INFORMATION: PHOSPHORYLATION, potential protein kinase C, Prosite match
NAME/KEY: MOD_RES
LOCATION: (147)..(147)
OTHER INFORMATION: PHOSPHORYLATION, potential caselase kinase II, Prosite match
NAME/KEY: MOD_RES
LOCATION: (194)..(194)
OTHER INFORMATION: PHOSPHORYLATION, potential protein kinase C, Prosite match
NAME/KEY: MOD_RES
LOCATION: (215)..(215)
OTHER INFORMATION: PHOSPHORYLATION, potential Tyrosine kinase site, Prosite match
NAME/KEY: MOD_RES
LOCATION: (221)..(221)
OTHER INFORMATION: SULFATATION, Prosite match
NAME/KEY: MOD_RES
LOCATION: (233)..(233)
OTHER INFORMATION: PHOSPHORYLATION, potential cAMP and cGMP dependant protein kinase
NAME/KEY: MOD_RES
LOCATION: (306)..(306)
OTHER INFORMATION: PHOSPHORYLATION, potential protein kinase C, Prosite match
NAME/KEY: HELIX
LOCATION: (310)..(330)
OTHER INFORMATION: Eisenberg, Schwarz, Komaromy, Wall identification method,
NAME/KEY: LIPID
LOCATION: (319)..(323)
OTHER INFORMATION: MYRISTATE, Prosite match
NAME/KEY: LIPID
LOCATION: (323)..(327)
OTHER INFORMATION: MYRISTATE, Prosite match
NAME/KEY: MOD_RES
LOCATION: (329)..(329)
OTHER INFORMATION: AMIDATION, Prosite match
NAME/KEY: HELIX
LOCATION: (333)..(353)
OTHER INFORMATION: Eisenberg, Schwarz, Komaromy, Wall identification method,
NAME/KEY: LIPID
LOCATION: (341)..(345)
OTHER INFORMATION: MYRISTATE, Prosite match
NAME/KEY: MOD_RES
LOCATION: (350)..(350)
OTHER INFORMATION: PHOSPHORYLATION, potential protein kinase C, Prosite match
US-09-901-484A-4
```

Query Match 97.2%; Score 1871; DB 23; Length 353;  
Best Local Similarity 100.0%; Pred. No. 1.1e-183;  
Matches 353; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY 12 MRYLLPSVLLGTAPTYVLAMGVRRLLSAFLPARFYQALDRLCYVQSNVLPFFENYTG 71
DB 1 MRYLLPSVLLGTAPTYVLAMGVRRLLSAFLPARFYQALDRLCYVQSNVLPFFENYTG 60
QY 72 VOILLGDLPRKNENIYYLANHOSTVDIMYADILAIQNALGHRVYLKGLKMLPLYGC 131
DB 61 VOILLGDLPRKNENIYYLANHOSTVDIMYADILAIQNALGHRVYLKGLKMLPLYGC 120
QY 132 YFAOHGGIYVRSKAFNEKEMRNKLOSYVDAGTPMYLVIFPEGTRYNPEQTKVLSASQAF 191
DB 121 YFAOHGGIYVRSKAFNEKEMRNKLOSYVDAGTPMYLVIFPEGTRYNPEQTKVLSASQAF 180
QY 192 AAQGLAVLKHVLTPIRKATIHVAFDCKKNYLDALYDVTVVYEGKDDGQRRSEPTMTBEFL 251
```

```
DB 181 AAQGLAVLKHVLTPIRKATIHVAFDCKKNYLDALYDVTVVYEGKDDGQRRSEPTMTBEFL 240
QY 252 CKECPKIHIDRIDDKDVPEOEHRMRWMLHEREIKDKMLIEFYESPDPERRRRPFQGS 311
DB 241 CKECPKIHIDRIDDKDVPEOEHRMRWMLHEREIKDKMLIEFYESPDPERRRRPFQGS 300
QY 312 VNSKLSIKKTLPSMLISGLTAGMLMTDAGRKLVTNWIYGTLLGCLMWTIKA 364
DB 301 VNSKLSIKKTLPSMLISGLTAGMLMTDAGRKLVTNWIYGTLLGCLMWTIKA 353

RESULT 15
US-09-915-582-59
Sequence 59, Application US/09915582
GENERAL INFORMATION:
APPLICANT: Rosen et al.
TITLE OR INVENTION: 17 Human Secreted Proteins
FILE REFERENCE: P5723P1
CURRENT APPLICATION NUMBER: US/09/915,582
PRIOR FILING DATE: 2001-07-27
PRIOR APPLICATION NUMBER: PCT/US01/01431
PRIOR FILING DATE: 2001-01-17
PRIOR APPLICATION NUMBER: 60/179,065
PRIOR FILING DATE: 2000-01-31
PRIOR APPLICATION NUMBER: 60/180,628
PRIOR FILING DATE: 2000-02-04
PRIOR APPLICATION NUMBER: 60/231,968
PRIOR FILING DATE: 2000-09-12
NUMBER OF SEQ ID NOS: 97
SOFTWARE: Patentl Ver. 2.0
SEQ ID NO 59
LENGTH: 353
TYPE: PRT
ORGANISM: Homo sapiens
US-09-915-582-59
```

Query Match 97.2%; Score 1871; DB 23; Length 353;  
Best Local Similarity 100.0%; Pred. No. 1.1e-183;  
Matches 353; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY 12 MRYLLPSVLLGTAPTYVLAMGVRRLLSAFLPARFYQALDRLCYVQSNVLPFFENYTG 71
DB 1 MRYLLPSVLLGTAPTYVLAMGVRRLLSAFLPARFYQALDRLCYVQSNVLPFFENYTG 60
QY 72 VOILLGDLPRKNENIYYLANHOSTVDIMYADILAIQNALGHRVYLKGLKMLPLYGC 131
DB 61 VOILLGDLPRKNENIYYLANHOSTVDIMYADILAIQNALGHRVYLKGLKMLPLYGC 120
QY 132 YFAOHGGIYVRSKAFNEKEMRNKLOSYVDAGTPMYLVIFPEGTRYNPEQTKVLSASQAF 191
DB 121 YFAOHGGIYVRSKAFNEKEMRNKLOSYVDAGTPMYLVIFPEGTRYNPEQTKVLSASQAF 180
QY 192 AAQGLAVLKHVLTPIRKATIHVAFDCKKNYLDALYDVTVVYEGKDDGQRRSEPTMTBEFL 251
DB 181 AAQGLAVLKHVLTPIRKATIHVAFDCKKNYLDALYDVTVVYEGKDDGQRRSEPTMTBEFL 240
QY 252 CKECPKIHIDRIDDKDVPEOEHRMRWMLHEREIKDKMLIEFYESPDPERRRRPFQGS 311
DB 241 CKECPKIHIDRIDDKDVPEOEHRMRWMLHEREIKDKMLIEFYESPDPERRRRPFQGS 300
QY 312 VNSKLSIKKTLPSMLISGLTAGMLMTDAGRKLVTNWIYGTLLGCLMWTIKA 364
DB 301 VNSKLSIKKTLPSMLISGLTAGMLMTDAGRKLVTNWIYGTLLGCLMWTIKA 353
```

Search completed: August 28, 2002, 11:14:11  
Job time: 490 sec

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 28, 2002, 11:15:41 ; Search time 84.08 Seconds  
(without alignments)  
1047.889 Million cell updates/sec

Title: US-09-853-526-5  
1924  
Sequence: 1 MLISLVHTYSMRFLPSVY.....YVNTWIVGTLLGLMTWTFKA 364

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 747981 seqs, 242050750 residues

Total number of hits satisfying chosen parameters: 747981

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

Pending\_Patents\_AA\_New:\*  
1: /cgn2\_6/ptodata/1/paa/US06\_NEW\_COMB.pep:\*  
2: /cgn2\_6/ptodata/1/paa/US06\_NEW\_COMB.pep:\*  
3: /cgn2\_6/ptodata/1/paa/US07\_NEW\_COMB.pep:\*  
4: /cgn2\_6/ptodata/1/paa/US08\_NEW\_COMB.pep:\*  
5: /cgn2\_6/ptodata/1/paa/US09\_NEW\_COMB.pep:\*  
6: /cgn2\_6/ptodata/1/paa/US10\_NEW\_COMB.pep:\*  
7: /cgn2\_6/ptodata/1/paa/US60\_NEW\_COMB.pep:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1924	100.0	364	US-10-184-648-63	Sequence 63, Appl
2	1868	97.1	353	US-09-629-469A-13028	Sequence 13028, A
3	1427	74.2	269	US-10-074-045-47	Sequence 47, Appl
4	390.5	20.3	375	US-09-935-625-11061	Sequence 11061, A
5	390.5	20.3	375	US-09-935-625-14197	Sequence 14197, A
6	390.5	20.3	375	US-09-935-625-14201	Sequence 14201, A
7	390.5	20.3	375	US-09-935-625-14254	Sequence 14254, A
8	390.5	20.3	375	US-09-935-625-23009	Sequence 23009, A
9	390.5	20.3	375	US-09-935-625-23025	Sequence 23025, A
10	387.5	20.1	351	US-09-935-625-11062	Sequence 11062, A
11	387.5	20.1	351	US-09-935-625-14198	Sequence 14198, A
12	387.5	20.1	351	US-09-935-625-14202	Sequence 14202, A
13	387.5	20.1	351	US-09-935-625-14255	Sequence 14255, A
14	387.5	20.1	351	US-09-935-625-23010	Sequence 23010, A
15	387.5	20.1	375	US-09-935-625-23026	Sequence 23026, A
16	387.5	20.1	375	US-09-935-625-11471	Sequence 11471, A
17	384.5	20.0	351	US-09-935-625-11472	Sequence 11472, A
18	381	19.8	343	US-09-935-625-11063	Sequence 11063, A
19	381	19.8	343	US-09-935-625-11473	Sequence 11473, A
20	381	19.8	343	US-09-935-625-14199	Sequence 14199, A
21	381	19.8	343	US-09-935-625-14203	Sequence 14203, A
22	381	19.8	343	US-09-935-625-14256	Sequence 14256, A
23	381	19.8	343	US-09-935-625-23011	Sequence 23011, A
24	381	19.8	343	US-09-935-625-23027	Sequence 23027, A
25	374.5	19.5	400	US-60-391-781-1157	Sequence 1157, Ap
26	305.5	15.9	414	US-10-121-062-314	Sequence 314, App

27	305.5	15.9	414	6	US-10-063-502-102	Sequence 102, App
28	305.5	15.9	414	6	US-10-063-510-102	Sequence 102, App
29	305.5	15.9	414	6	US-10-063-512-102	Sequence 102, App
30	305.5	15.9	414	6	US-10-063-513-102	Sequence 102, App
31	305.5	15.9	414	6	US-10-063-514-102	Sequence 102, App
32	305.5	15.9	414	6	US-10-063-515-102	Sequence 102, App
33	305.5	15.9	414	6	US-10-063-516-102	Sequence 102, App
34	305.5	15.9	414	6	US-10-063-517-102	Sequence 102, App
35	305.5	15.9	414	6	US-10-063-518-102	Sequence 102, App
36	305.5	15.9	414	6	US-10-063-519-102	Sequence 102, App
37	305.5	15.9	414	6	US-10-063-520-102	Sequence 102, App
38	305.5	15.9	414	6	US-10-063-521-102	Sequence 102, App
39	305.5	15.9	414	6	US-10-063-522-102	Sequence 102, App
40	305.5	15.9	414	6	US-10-063-523-102	Sequence 102, App
41	305.5	15.9	414	6	US-10-063-524-102	Sequence 102, App
42	305.5	15.9	414	6	US-10-063-525-102	Sequence 102, App
43	305.5	15.9	414	6	US-10-063-526-102	Sequence 102, App
44	305.5	15.9	414	6	US-10-063-527-102	Sequence 102, App
45	305.5	15.9	414	6	US-10-063-528-102	Sequence 102, App
					US-10-063-529-102	Sequence 102, App

#### ALIGNMENTS

RESULT 1  
US-10-184-648-63  
Sequence 63, Application US/10184648  
GENERAL INFORMATION:  
APPLICANT: Meyers, Rachel E.  
APPLICANT: Williamson, Mark  
APPLICANT: Tsai, Fong-Ying  
APPLICANT: Hunter, John J.  
APPLICANT: Macbeth, Kyle J.  
APPLICANT: Rudolph-Owen, Laura A.  
APPLICANT: Leiby, Kevin R.  
APPLICANT: Kapeller-Libermann, Rosana  
APPLICANT: Olandt, Peter J.  
TITLE OF INVENTION: NOVEL HUMAN TRANSFERASE FAMILY MEMBERS AND USES THEREOF  
FILE REFERENCE: 10448-192001  
CURRENT APPLICATION NUMBER: US/10/184, 648  
CURRENT FILING DATE: 2002-06-27  
PRIOR APPLICATION NUMBER: US 09/815, 028  
PRIOR FILING DATE: 2001-03-22  
PRIOR APPLICATION NUMBER: PCT/US01/09358  
PRIOR FILING DATE: 2001-03-22  
PRIOR APPLICATION NUMBER: US 60/191, 964  
PRIOR FILING DATE: 2000-03-24  
PRIOR APPLICATION NUMBER: US 09/801, 220  
PRIOR FILING DATE: 2001-03-07  
PRIOR APPLICATION NUMBER: PCT/US01/07269  
PRIOR FILING DATE: 2001-03-07  
PRIOR APPLICATION NUMBER: US 60/187, 456  
PRIOR FILING DATE: 2000-03-07  
PRIOR APPLICATION NUMBER: US 09/816, 714  
PRIOR FILING DATE: 2001-03-23  
PRIOR APPLICATION NUMBER: PCT/US01/09468  
PRIOR FILING DATE: 2001-03-23  
PRIOR APPLICATION NUMBER: US 60/191, 865  
PRIOR FILING DATE: 2000-03-24  
PRIOR APPLICATION NUMBER: US 09/844, 948  
PRIOR FILING DATE: 2001-04-27  
PRIOR APPLICATION NUMBER: PCT/US01/13805  
PRIOR FILING DATE: 2001-04-27  
PRIOR APPLICATION NUMBER: US 60/200, 604  
PRIOR FILING DATE: 2000-04-28  
PRIOR APPLICATION NUMBER: US 09/861, 164  
PRIOR FILING DATE: 2001-05-18  
PRIOR APPLICATION NUMBER: PCT/US01/16292  
PRIOR FILING DATE: 2001-05-18  
PRIOR APPLICATION NUMBER: US 60/205, 408  
PRIOR FILING DATE: 2000-05-19  
PRIOR APPLICATION NUMBER: US 09/883, 060  
PRIOR FILING DATE: 2001-06-15

```
? PRIOR APPLICATION NUMBER: PCT/US01/19138
? PRIOR FILING DATE: 2001-06-15
? PRIOR APPLICATION NUMBER: US 60/212,079
? PRIOR FILING DATE: 2000-06-15
? PRIOR APPLICATION NUMBER: US 09/962,678
? PRIOR FILING DATE: 2001-09-25
? PRIOR APPLICATION NUMBER: PCT/US01/29963
? PRIOR FILING DATE: 2001-09-25
? PRIOR APPLICATION NUMBER: US 60/235,044
? PRIOR FILING DATE: 2000-09-25
? PRIOR APPLICATION NUMBER: US 09/973,457
? PRIOR FILING DATE: 2001-10-09
? PRIOR APPLICATION NUMBER: US 60/238,849
? PRIOR FILING DATE: 2000-10-06
? PRIOR APPLICATION NUMBER: US 10/072,285
? PRIOR FILING DATE: 2002-02-08
? PRIOR APPLICATION NUMBER: PCT/US02/03736
? PRIOR FILING DATE: 2002-02-08
? PRIOR APPLICATION NUMBER: US 60/267,494
? PRIOR FILING DATE: 2001-02-08
? PRIOR APPLICATION NUMBER: US 09/817,910
? PRIOR FILING DATE: 2001-03-26
? PRIOR APPLICATION NUMBER: PCT/US01/09633
? PRIOR FILING DATE: 2001-03-26
? PRIOR APPLICATION NUMBER: US 60/192,092
? PRIOR FILING DATE: 2000-03-24
? PRIOR APPLICATION NUMBER: US 09/842,528
? PRIOR FILING DATE: 2001-04-25
? PRIOR APPLICATION NUMBER: PCT/US01/40607
? PRIOR FILING DATE: 2001-04-25
? PRIOR APPLICATION NUMBER: US 60/199,500
? PRIOR FILING DATE: 2000-04-25
? PRIOR APPLICATION NUMBER: US 09/882,836
? PRIOR FILING DATE: 2001-06-15
? PRIOR APPLICATION NUMBER: PCT/US01/19543
? PRIOR FILING DATE: 2001-06-15
? PRIOR APPLICATION NUMBER: US 60/211,730
? PRIOR FILING DATE: 2000-06-15
? PRIOR APPLICATION NUMBER: US 09/882,872
? PRIOR FILING DATE: 2001-06-15
? PRIOR APPLICATION NUMBER: PCT/US01/19153
? PRIOR FILING DATE: 2001-06-15
? PRIOR APPLICATION NUMBER: US 60/212,077
? PRIOR FILING DATE: 2000-06-15
? NUMBER OF SEQ ID NOS: 90
? SOFTWARE: FastSeq for Windows Version 4.0
? SEQ ID NO 63
? LENGTH: 364
? TYPE: PRT
? ORGANISM: Homo sapiens
US-10-184-648-63
```

```
Query Match 100.0%; Score 1924; DB 6; Length 364;
Best Local Similarity 100.0%; Pred. No. 1.5e-171;
Matches 364; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MLTSLVHTYSMRRLPSVVLGTAPTYVLAMGWRLSAPLPARFYALDDRLCYVQOS 60
DB 1 MLTSLVHTYSMRRLPSVVLGTAPTYVLAMGWRLSAPLPARFYALDDRLCYVQOS 60
QY 61 MLEFFENYTGVOILLYGDLPKRNENIYLANHOSTVDMIVADILAIKRONALGHVRYVLK 120
DB 61 MLEFFENYTGVOILLYGDLPKRNENIYLANHOSTVDMIVADILAIKRONALGHVRYVLK 120
QY 121 EGLKWLPLYGIFYAQHGGIYVRSKAKNEKMRNKLQSYVDAGTPMYLVIFPEGTRYNPE 180
DB 121 EGLKWLPLYGIFYAQHGGIYVRSKAKNEKMRNKLQSYVDAGTPMYLVIFPEGTRYNPE 180
QY 181 OTKVLASQAFPAORGLAVLKHVLTLPRIKATHVAFDCMKNTLDALYDVTVYVEGKDDGGO 240
DB 181 OTKVLASQAFPAORGLAVLKHVLTLPRIKATHVAFDCMKNTLDALYDVTVYVEGKDDGGO 240
```

```
QY 241 RRESPTMTFLCKECPKIHIDRIDKKVPEEGEHMRWLHERFEIKDKMLIEFYSPD 300
DB 241 RRESPTMTFLCKECPKIHIDRIDKKVPEEGEHMRWLHERFEIKDKMLIEFYSPD 300
QY 301 PERKRRPGRKSVNSKLSIKTKLPMSLLISGLTAGMLMTDAGRKYVYMWYIGTLGLMW 360
DB 301 PERKRRPGRKSVNSKLSIKTKLPMSLLISGLTAGMLMTDAGRKYVYMWYIGTLGLMW 360
QY 361 TTKA 364
DB 361 TTKA 364
```

```
RESULT 2
US-09-629-469A-13028
? Sequence 13028, Application us/09629469A
? GENERAL INFORMATION:
? APPLICANT: OTA, TOSHIO
? APPLICANT: ISOGAI, TAKAO
? APPLICANT: NISHIKAWA, TETSUO
? APPLICANT: HAYASHI, KOJI
? APPLICANT: SATTO, KAORU
? APPLICANT: YAMAMOTO, JUNICHI
? APPLICANT: ISHII, SHIZUKO
? APPLICANT: SUGIYAMA, TOMOYASU
? APPLICANT: WAKAMATSU, AI
? APPLICANT: NAGAI, KEIICHI
? APPLICANT: OTSUKI, TETSUJI
? TITLE OF INVENTION: PRIMERS FOR SYNTHESIZING FULL-LENGTH CDNA AND THEIR USE
? FILE REFERENCE: 084335/0123
? CURRENT APPLICATION NUMBER: US/09/629,469A
? PRIOR FILING DATE: 2000-07-28
? PRIOR APPLICATION NUMBER: JP 1999-248036
? PRIOR FILING DATE: 1999-07-29
? PRIOR APPLICATION NUMBER: JP 1999-300253
? PRIOR FILING DATE: 1999-08-27
? PRIOR APPLICATION NUMBER: JP 2000-118776
? PRIOR FILING DATE: 2000-01-11
? PRIOR APPLICATION NUMBER: JP 2000-183767
? PRIOR FILING DATE: 2000-05-02
? PRIOR APPLICATION NUMBER: JP 2000-241899
? PRIOR FILING DATE: 2000-06-09
? PRIOR APPLICATION NUMBER: 60/159,590
? PRIOR FILING DATE: 1999-10-18
? PRIOR APPLICATION NUMBER: 60/183,322
? PRIOR FILING DATE: 2000-02-17
? NUMBER OF SEQ ID NOS: 19025
? SOFTWARE: PatentIn Ver. 2.1
? SEQ ID NO 13028
? LENGTH: 353
? TYPE: PRT
? ORGANISM: Homo sapiens
US-09-629-469A-13028
```

```
Query Match 97.1%; Score 1866; DB 5; Length 353;
Best Local Similarity 99.7%; Pred. No. 2.5e-166;
Matches 352; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 12 MRYLLPSVVLGTAPTYVLAMGWRLSAPLPARFYALDDRLCYVQOSWVLPFFENYTG 71
DB 1 MRYLLPSVVLGTAPTYVLAMGWRLSAPLPARFYALDDRLCYVQOSWVLPFFENYTG 71
QY 72 VOILLYGDLPKRNENIYLANHOSTVDMIVADILAIKRONALGHVRYVLKGLKMLPLYGC 131
DB 72 VOILLYGDLPKRNENIYLANHOSTVDMIVADILAIKRONALGHVRYVLKGLKMLPLYGC 131
QY 132 YFAOHGGIYVRSKAKNEKMRNKLQSYVDAGTPMYLVIFPEGTRYNPEQTKVLSASQAF 191
DB 132 YFAOHGGIYVRSKAKNEKMRNKLQSYVDAGTPMYLVIFPEGTRYNPEQTKVLSASQAF 191
QY 191 YFAOHGGIYVRSKAKNEKMRNKLQSYVDAGTPMYLVIFPEGTRYNPEQTKVLSASQAF 180
DB 191 YFAOHGGIYVRSKAKNEKMRNKLQSYVDAGTPMYLVIFPEGTRYNPEQTKVLSASQAF 180
QY 192 AAORGLAVLKHVLTLPRIKATHVAFDCMKNTLDALYDVTVYVEGKDDGGRRESPTMTFEL 251
DB 192 AAORGLAVLKHVLTLPRIKATHVAFDCMKNTLDALYDVTVYVEGKDDGGRRESPTMTFEL 251
```

```
Db 181 AAORGLAVLKHVLTPTRIKATHVAFDCMKNYLDAIDVTYVVEEGKDDGGORRESPTMTFEL 240
QY 252 CKECPKIHIDIRIDKDDVPEOEHRMRMLHEFEIKDKMLIEFYSPDERRRKRRPGKS 311
Db 241 CKECPKIHIDIRIDKDDVPEOEHRMRMLHEFEIKDKMLIEFYSPDERRRKRRPGKS 300
QY 312 VNSKLSIKKTLPSSMLILSGLTAGMLMTDAGRKLYVTWYITGLGCLMWYIKA 364
Db 301 VNSKLSIKKTLPSSMLILSGLTAGMLMTDAGRKLYVTWYITGLGCLMWYIKA 353

RESULT 3
US-10-074-045-47
; Sequence 47, Application US/10074045
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PT2121
; CURRENT APPLICATION NUMBER: US/10/074,045
; CURRENT FILING DATE: 2002-02-14
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 74
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 47
; LENGTH: 269
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-074-045-47
```

```
Query Match 74.2%; Score 1427; DB 6; Length 269;
Best Local Similarity 100.0%; Pred. No. 3.3e-125;
Matches 269; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 96 TYDMYADLAIARONALGHVRYVLEKGLKWLPLGYCYFAOHGSIYVRSKAFNEKEMRNK 155
Db 1 TYDMYADLAIARONALGHVRYVLEKGLKWLPLGYCYFAOHGSIYVRSKAFNEKEMRNK 60
QY 156 LOSYVAGTPTMYLTPEEGTRVNEBOTKYLASQAFAGLAVLKHVLTPTRIKATHVAF 215
Db 61 LOSYVAGTPTMYLTPEEGTRVNEBOTKYLASQAFAGLAVLKHVLTPTRIKATHVAF 120
QY 216 DCMKNYLDAIDVTYVVEEGKDDGGORRESPTMTFELCKECPKIHIDIRIDKDDVPEOE 275
Db 121 DCMKNYLDAIDVTYVVEEGKDDGGORRESPTMTFELCKECPKIHIDIRIDKDDVPEOE 180
QY 276 HMRMLHEFEIKDKMLIEFYSPDERRRKRRPGKS VNSKLSIKKTLPSSMLILSGLTAGM 335
Db 181 HMRMLHEFEIKDKMLIEFYSPDERRRKRRPGKS VNSKLSIKKTLPSSMLILSGLTAGM 240
QY 336 LMTDAGRKLYVTWYITGLGCLMWYIKA 364
Db 241 LMTDAGRKLYVTWYITGLGCLMWYIKA 269
```

```
RESULT 4
US-09-935-625-11061
; Sequence 11061, Application US/09935625
; GENERAL INFORMATION:
; APPLICANT: N. ALEXANDROV et al.
; TITLE OF INVENTION: POLYNUCLEOTIDES, POLYPEPTIDES, CELLS, AND METHODS THEREOF CAPABLE
; FILE REFERENCE: 2750-1481P
; CURRENT APPLICATION NUMBER: US/09/935,625
; CURRENT FILING DATE: 2001-08-24
; NUMBER OF SEQ ID NOS: 33136
; SEQ ID NO 11061
; LENGTH: 375
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; NAME/KEY: peptide
; LOCATION: 1..375
```

```
; OTHER INFORMATION: Ceres Seq. ID no. 1394370
US-09-935-625-11061
```

```
Query Match 20.3%; Score 390.5; DB 5; Length 375;
Best Local Similarity 26.9%; Pred. No. 5.1e-28;
Matches 101; Conservative 78; Mismatches 128; Indels 69; Gaps 9;
```

```
QY 12 MRYLPSVVLGTAPTYVLAMG-----VWRLSAPLPARYQALDRLYCYOSMVLFEFF 66
Db 18 LRGIICLMLVLTSTAFMMLIFWGLSAVVLRLFS---IRYSRCVSEFFSGMLALMPLEF 73
QY 67 ENTGVQIILYGDLPKKNENIYLANHOSVDMIVADILAIARONALGHVRYVLEKGLKWL 126
Db 74 EKINKTKVITSGDKVPCEDEVLLIANHRTVDMMYFMDLAKRGQGINIKYIVLKSLSMKL 133
QY 127 PLVGYFAOHGSIYVRSKAFNEKEMRNKLOSYVDACTPTMYLVIPEEGTRVNEBOTKYL 186
Db 134 PLFGMAFHLFEFTIPVERWEDEVANLRLQIVSFPKPRDALMLALFEGTDTYEAQCO--- 190
QY 187 ASQAFPAORGLAVLKHVLTPTRIKATHVAFDCMKNYLDAIDVTYVVEEGKDDGGORRESPT 246
Db 191 RSKKFAENGLPILNNVLTPTKGFVSCLOELSCSDAVYDTYIGKTR----- 239
QY 247 MTEFLCKECP-----KIHIDIRIDKDDVPEOEHRMRMLHEFEIKDKMLIE 294
Db 240 -----CPSFLDNVYGIPSEVHIHRIINLTQIPNOKDINAMLMNTFOLKDQLND 291
QY 295 FYESPDERRRKRRPGKS VNSKLSIKKTLPSSMLILSGLT-----AGMLMTDAGRKLY 345
Db 292 FYSN-----GHPNNECTEKEFTMKKYLINCLAVIATTTICTHLPFFSSIMWF----RIV 341
QY 346 VNTWYITGLGCLMWY 361
Db 342 VS-----LACYLT 350
```

```
RESULT 5
US-09-935-625-14197
; Sequence 14197, Application US/09935625
; GENERAL INFORMATION:
; APPLICANT: N. ALEXANDROV et al.
; TITLE OF INVENTION: POLYNUCLEOTIDES, POLYPEPTIDES, CELLS, AND METHODS THEREOF CAPA
; FILE REFERENCE: 2750-1481P
; CURRENT APPLICATION NUMBER: US/09/935,625
; CURRENT FILING DATE: 2001-08-24
; NUMBER OF SEQ ID NOS: 33136
; SEQ ID NO 14197
; LENGTH: 375
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; NAME/KEY: peptide
; LOCATION: 1..375
; OTHER INFORMATION: Ceres Seq. ID no. 3441314
US-09-935-625-14197
```

```
Query Match 20.3%; Score 390.5; DB 5; Length 375;
Best Local Similarity 26.9%; Pred. No. 5.1e-28;
Matches 101; Conservative 78; Mismatches 128; Indels 69; Gaps 9;

QY 12 MRYLPSVVLGTAPTYVLAMG-----VWRLSAPLPARYQALDRLYCYOSMVLFEFF 66
Db 18 LRGIICLMLVLTSTAFMMLIFWGLSAVVLRLFS---IRYSRCVSEFFSGMLALMPLEF 73
QY 67 ENTGVQIILYGDLPKKNENIYLANHOSVDMIVADILAIARONALGHVRYVLEKGLKWL 126
Db 74 EKINKTKVITSGDKVPCEDEVLLIANHRTVDMMYFMDLAKRGQGINIKYIVLKSLSMKL 133
QY 127 PLVGYFAOHGSIYVRSKAFNEKEMRNKLOSYVDACTPTMYLVIPEEGTRVNEBOTKYL 186
Db 134 PLFGMAFHLFEFTIPVERWEDEVANLRLQIVSFPKPRDALMLALFEGTDTYEAQCO--- 190
```





```
; LOCATION: 1..375
; OTHER INFORMATION: Ceres Seq. ID no. 3441318
US-09-935-625-23009

Query Match
Best Local Similarity 20.3%; Score 390.5; DB 5; Length 375;
Matches 101; Conservative 78; Mismatches 128; Indels 69; Gaps 9;

OY 12 MRLTLPVLLGTAFTYVLWG-----VWRLLSAFLPARFYQALDRLCYQOSMVLFEFF 66
DB 18 LRGIILMLVSTAFMMLIFWGFLSAVVLRLFS---IRYSRKCVSFFGSMWALMPFLF 73
OY 67 ENTGVOILLGDLPRKKNKIITLANHSTVDWIVADILAIRONALGHVRYVLEKGLKWL 126
DB 74 EKINRKRVIFSGDKVCEDEVLLIANHRTVDMMYFMDLIRKGOINITYLVKSSLMKL 133
OY 127 PLYGCFYAOHGIIYVRSKAFNEKEMRNKLQSYVDAGTPLYVLFPEGTRYNDEQTVLS 186
DB 134 PLFGMAFHLFEFIPVERRWEVDANLRQIVSFPKPRDALMLALFPEGTDYTEAKCO--- 190
OY 187 ASQAFPAORGLAVLKHLVTRIKATHVAFDCMKNYLDATIDYVYVEGKDDGGORRESPT 246
DB 191 RSKRFAENGLPILNNVLPRTKGFSCLQELSCSDAVVDVITGYKTR----- 239
OY 247 MTEFLCKECP-----KIHIIHIDRIKDKVPEDEQEHMRMLHERFEIKDKMLIE 294
DB 240 -----CPSFLDNVYGIPESEVHHIRKINLTQIPNOEKDINAMLMNTFOLKDLND 291
OY 295 FYESPDERRRKRPKGSVNSKLSIKTLPMSLILSGLT-----AGMLMTDAGRKLX 345
DB 292 FYSN-----GHFPENGTEKEFNTKVKYLINCIAVIAFTTICTHLPFFSSMIWF---RIY 341
OY 346 VNTWYIGTLGCLMWT 361
DB 342 VS-----LACVYLT 350

RESULT 9
US-09-935-625-23025
; Sequence 23025, Application US/09935625
; GENERAL INFORMATION:
; APPLICANT: N. ALEXANDROV et al.
; TITLE OF INVENTION: POLYNUCLEOTIDES, POLYPEPTIDES, CELLS, AND METHODS THEREOF CAPABLE
; FILE REFERENCE: 2750-1481P
; CURRENT APPLICATION NUMBER: US/09/935,625
; CURRENT FILING DATE: 2001-08-24
; NUMBER OF SEQ ID NOS: 33136
; SEQ ID NO 23025
; LENGTH: 375
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; NAME/KEY: peptide
; LOCATION: 1..375
; OTHER INFORMATION: Ceres Seq. ID no. 3443248
US-09-935-625-23025

Query Match
Best Local Similarity 20.3%; Score 390.5; DB 5; Length 375;
Matches 101; Conservative 78; Mismatches 128; Indels 69; Gaps 9;
```

```
DB 134 PLFGMAFHLFEFIPVERRWEVDANLRQIVSFPKPRDALMLALFPEGTDYTEAKCO--- 190
OY 187 ASQAFPAORGLAVLKHLVTRIKATHVAFDCMKNYLDATIDYVYVEGKDDGGORRESPT 246
DB 191 RSKRFAENGLPILNNVLPRTKGFSCLQELSCSDAVVDVITGYKTR----- 239
OY 247 MTEFLCKECP-----KIHIIHIDRIKDKVPEDEQEHMRMLHERFEIKDKMLIE 294
DB 240 -----CPSFLDNVYGIPESEVHHIRKINLTQIPNOEKDINAMLMNTFOLKDLND 291
OY 295 FYESPDERRRKRPKGSVNSKLSIKTLPMSLILSGLT-----AGMLMTDAGRKLX 345
DB 292 FYSN-----GHFPENGTEKEFNTKVKYLINCIAVIAFTTICTHLPFFSSMIWF---RIY 341
OY 346 VNTWYIGTLGCLMWT 361
DB 342 VS-----LACVYLT 350

RESULT 10
US-09-935-625-11062
; Sequence 11062, Application US/09935625
; GENERAL INFORMATION:
; APPLICANT: N. ALEXANDROV et al.
; TITLE OF INVENTION: POLYNUCLEOTIDES, POLYPEPTIDES, CELLS, AND METHODS THEREOF CAPA
; FILE REFERENCE: 2750-1481P
; CURRENT APPLICATION NUMBER: US/09/935,625
; CURRENT FILING DATE: 2001-08-24
; NUMBER OF SEQ ID NOS: 33136
; SEQ ID NO 11062
; LENGTH: 351
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; NAME/KEY: peptide
; LOCATION: 1..351
; OTHER INFORMATION: Ceres Seq. ID no. 1394371
US-09-935-625-11062

Query Match
Best Local Similarity 20.1%; Score 387.5; DB 5; Length 351;
Matches 100; Conservative 75; Mismatches 125; Indels 69; Gaps 9;

OY 19 VLLGTAFTYVLWG-----VWRLLSAFLPARFYQALDRLCYQOSMVLFEFFENTYGVQ 73
DB 1 MVLVSTAFMMLIFWGFLSAVVLRLFS---IRYSRKCVSFFGSMWALMPFLFEKINKTK 56
OY 74 ILLYGLDLPKKNKIITLANHSTVDWIVADILAIRONALGHVRYVLEKGLKMLPLGYCYF 133
DB 57 VIFSGDKVPCEDRVLLIANHRTVDMMYFMDLIRKGOINITYLVKSSLMKLPLFSGWAF 116
OY 134 AQHGGIYVRSKAFNEKEMRNKLQSYVDAGTPLYVLFPEGTRYNDEQTVLSASQFAA 193
DB 117 HLEFIPVERRWEVDANLRQIVSFPKPRDALMLALFPEGTDYTEAKCO---RSKFAA 173
OY 194 QRGVLAVLKHLVTRIKATHVAFDCMKNYLDATIDYVYVEGKDDGGORRESPTMTEFLCK 253
DB 174 ENGLPILNNVLPRTKGFSCLQELSCSDAVVDVITGYKTR----- 215
OY 254 ECP-----KIHIIHIDRIKDKVPEDEQEHMRMLHERFEIKDKMLIEFYESPDP 301
DB 216 -CPSFLDNVYGIPESEVHHIRKINLTQIPNOEKDINAMLMNTFOLKDLNDFYSN-- 271
OY 302 ERRKRFPKGSVNSKLSIKTLPMSLILSGLT-----AGMLMTDAGRKLXVNTWYIG 352
DB 272 -----GHFPENGTEKEFNTKVKYLINCIAVIAFTTICTHLPFFSSMIWF---RIYVS----- 319
OY 353 TLLGCLMWT 361
DB 320 --LACVYLT 326
```



[illegible]

```

RESULT 14
US-09-935-625-23010
: Sequence 23010, Application US/09935625
: GENERAL INFORMATION:
: APPLICANT: N. ALEXANDROV et al.
: TITLE OF INVENTION: POLYNUCLEOTIDES, POLYPEPTIDES, CELLS, AND METHODS THEREOF CAPABLE
: OF INDUCING A MODULATING EFFECT ON A CELL
: FILE REFERENCE: 2750-1481P
: CURRENT APPLICATION NUMBER: US/09/935,625
: CURRENT FILING DATE: 2001-08-24
: NUMBER OF SEQ ID NOS: 33136
: SEQ ID NO 23010
: LENGTH: 351
: TYPE: PRT
: ORGANISM: Arabidopsis thaliana
: FEATURE:
: NAME/KEY: peptide
: LOCATION: 1..351
: OTHER INFORMATION: Ceres Seq. ID no. 3441319
: US-09-935-625-23010

```

Query Match	20.1%;	Score 387.5;	DB 5;	Length 351;
Best Local Similarity	27.1%;	Pred. No. 8.9e-28;		
Matches 100;	Conservative 75;	Mismatches 125;	Indels 69;	Gaps 9

```

QY 19 VLLGTAPPTVYLAWG-----VWRLISFPLPARYQALDDRLXCVYOSMYLFEFFENYTAGV 73
   11 : 11 : 11 : 11 : 11 : 11 : 11 : 11 : 11 : 11 : 11 : 11 : 11 : 11 : 11 :
Db 1 MVLVSTAFMMILFENGFLSAYVLRFS----IRSRKCVSFFFSQWALMPLFEKINKRK 56

QY 74 ILLYGLPKNKENITITLANHOSYVDWIVADILAIRONALGHVRYVLEKGLKMLPLYGYCF 133
   11 : 11 : 11 : 11 : 11 : 11 : 11 : 11 : 11 : 11 : 11 : 11 : 11 : 11 :
Db 57 VIFSGDVPCEDRVLLTANHRTEVDWYFDLAIKRGQIGNIKYVLSKSLMKPLRGMAF 116

QY 134 AQHSGIVYVKSASAKNEKEMKNKLOSVDYDAPMLVLFPEGTRYNDPQFVLASAOAPA 193
   11 : 11 : 11 : 11 : 11 : 11 : 11 : 11 : 11 : 11 : 11 : 11 : 11 : 11 :
Db 117 HLEFIEYERKMEVDENLQIYVSFDPDPDALMLALFEGTIDYTEAKCO--RSKKFAA 173

QY 194 QRLGVALKHVLTPIRIKATHAVFDMKNKYLDIADYTVVYEGKIDGGORRSPMTPEFLCK 253
   11 : 11 : 11 : 11 : 11 : 11 : 11 : 11 : 11 : 11 : 11 : 11 : 11 :
Db 174 ENGLPIINLVLLPRTKGFSCLQELSCSLDPAVDVLTIGYKTR----- 215

QY 254 ECP-----KHHIDRIDDKDQVDEEBOBHKRMVLEHFEFLKDKMLIEFTSPDP 301
   11 : 11 : 11 : 11 : 11 : 11 : 11 : 11 : 11 : 11 : 11 : 11 : 11 :
Db 216 -CPSEFLDNVYGIEPSEVHHIRIRINLTQIPOEKDINAMLMNFOLKDLNDPVEN--- 271

QY 302 ERRKRPFGKSVNSKLTSLKTLPSMLISGLT-----AGMLMTDAGKELVNNWIVG 352
   11 : 11 : 11 : 11 : 11 : 11 : 11 : 11 : 11 : 11 : 11 : 11 : 11 :
Db 272 ---GHFNEGTEKEKFNKTKVILNCLAVIATFTTCTHLTFESSMIWF-----RIVS----- 319

QY 353 TLLGLAMWT 361
   11 : 11 : 11 : 11 : 11 : 11 : 11 : 11 : 11 : 11 : 11 : 11 : 11 :

```

```

Db          320 --LACCYLT 326

RESULT      15
US-09-935-625-23026
: Sequence 23026, Application US/09935625
GENERAL INFORMATION:
APPLICANT: N. ALEXANDROV et al
TITLE OF INVENTION: POLYNUCLEOTIDES, POLYPEPTIDES, CELLS, AND METHODS THEREOF CAPA
TITLE OF INVENTION: POLYNUCLEOTIDES, POLYPEPTIDES, CELLS, AND METHODS THEREOF CAPA
FILE REFERENCE: 2750-1481P
CURRENT APPLICATION NUMBER: US/09/935,625
CURRENT FILING DATE: 2001-08-24
NUMBER OF SEQ ID NOS: 33136
SEQ ID NO 23026
LENGTH: 351
TYPE: PR1
ORGANISM: Arabidopsis thaliana
FEATURE:
NAME/KEY: peptide
LOCATION: 1..351
OTHER INFORMATION: Ceres Seq. ID no. 3443249
US-09-935-625-23026

```

Query Match	20.1%;	Score 387.5;	DB 5;	Length 351;
Best Local Similarity	27.1%;	Pred. No. 8.9e-28;		
Matches 100;	Conservative 75;	Mismatches 125;	Indels 69;	Gaps 9

```

QY      19  VLLIGTAPTVLWLG-----VMRLSAPFLPAREFOALDDRLVCYOSMVEFFENYTGVO 73
Db      1  MVLVSTAFMWMLIFMGFLSTAVLRLFS----IRSRKCVSFFPSMLALMPELEKINKRK 56

QY      74  ILLYGLPPKKENITYLANHOSQVWDIVADILAIRONALGHVRYVLKGLKMLPLCYGF 133
Db      57  VIESGDVCPEDRVLILANHRTVEVDWMYFMDLIRKQOLINIKYVLSLMLKPLGMAF 116

QY      134 AQHSGIYVKKSAKFNEMENMKLOSQVADATPVYLVFPEGSTRYNNPEOTVLSASQAFNA 193
Db      117 HLEEFIVERMERWEDEANLQIVSSFDPPDALMLALFPEGTYTEAKCO--RSKKFPA 173

QY      194 ORGLAVLKHVTLTPRIKATNHAFOCMKNKYDAIVDVYVYEGKDGGCGRRESPTMFECLK 253
Db      174 ENGLPIANVLLPRTKGFVSCLOELSCSLDAVYDVITIGYTR----- 215

QY      254 ECP-----KIHIDRIDKKDVEBOEHMRKWLHERFEIKDKMLIEFYESPDP 301
Db      216 -CPSEFDLNVYGIPESEVHHIRIRINLTQIPNOEKDINAMLMNFPOLKQDLNDFYSN--- 271

QY      302 ERRKRPEGSVNSKLSIKTKLPSMLIISGLT-----AGMLMDAGKRLVYNTWITYG 352
Db      272 ---GHFNEGETEKEFNFKKYLINCLAVIAFTTICTHILTEPSSIMWF---RIVVS----- 319

QY      353 TLIGGLMVT 361
Db      320 -LACYPLT 326

```

Search completed: August 28, 2002, 11:15:42  
Job time: 556 sec



GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 28, 2002, 11:08:57 ; Search time 39.13 Seconds

(without alignments)  
893.854 Million cell updates/sec

Title: US-09-853-526-5

Perfect score: 1924

Sequence: 1 MLSTLVHTYSMRYLPSV.....YVNTWIVGTLLGLWVTIKA 364

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283138 seqs, 96089334 residues

Total number of hits satisfying chosen parameters: 283138

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : PIR\_71.\*

1: pir1.\*

2: pir2.\*

3: pir3.\*

4: pir4.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	470	24.4	918	2	T34057
2	398.5	20.7	393	2	B96780
3	328	17.0	374	2	S52645
4	307	16.0	376	2	S60478
5	263	13.7	376	2	D96550
6	249.5	13.0	350	2	T40466
7	242.5	12.6	344	2	T31913
8	239	12.4	311	2	T07936
9	238.5	12.4	310	2	T06755
10	231.5	12.0	397	2	S45900
11	221.5	11.5	396	2	S54641
12	213.5	11.1	523	2	T25998
13	212	11.0	439	2	T22689
14	184	9.6	435	2	B83587
15	182	9.5	391	2	T15366
16	164.5	8.5	310	2	S40808
17	164.5	8.5	310	2	A91227
18	164.5	8.5	310	2	G86073
19	153	8.0	304	2	B83541
20	145.5	7.6	285	2	T15252
21	136.5	7.1	302	2	A10950
22	131.5	6.8	362	2	T20608
23	130	6.8	294	2	D82371
24	124	6.4	303	2	A48600
25	118	6.1	281	2	S60477
26	116	6.0	266	2	S73863
27	114.5	6.0	247	2	G72223
28	112	5.8	258	2	G83101
29	111.5	5.8	377	1	A44216

30	111.5	5.8	488	1	Q0EB8S	alkaline exonuclease
31	109.5	5.7	488	2	T44030	alkaline exonuclease
32	109.5	5.7	488	2	T44215	alkaline exonuclease
33	109	5.7	240	2	D64089	1-acylglycerol-3-P
34	108	5.6	261	2	C82067	probable 1-acylglycerol
35	104	5.4	240	2	D64688	1-acylglycerol-3-P
36	104	5.4	267	2	AC3499	1-acylglycerol-3-P
37	103	5.4	243	2	AC0083	1-acylglycerol-3-P
38	103	5.4	266	2	B98283	hypothetical prote
39	103	5.4	266	2	AG3000	1-acyl-sn-glycerol
40	102	5.3	212	2	G85357	hypothetical prote
41	102	5.3	1551	2	T18941	hypothetical prote
42	101	5.2	262	2	T22599	1-acylglycerol-3-P
43	101	5.2	323	2	C83940	sugar ABC transport
44	99.5	5.1	237	2	B71827	probable 1-acylglycerol
45	99	5.1	225	2	B45582	probable 1-acylglycerol

ALIGNMENTS

RESULT	1	Score 470; DB 2; Length 918;
T34057	hypothetical protein F28B3.5 - Caenorhabditis elegans	
C:Species: Caenorhabditis elegans		
C:Date: 29-Oct-1999	#sequence_revision 29-Oct-1999	#text_change 29-Oct-1999
C:Accession: T34057		
R:Geisel, C.; Kramer, J.; Smith, A.		
A:Submitted to the EMBL Data Library, May 1997		
A:Description: The sequence of C. elegans cosmid F28B3.		
A:Reference number: 221469		
A:Accession: T34057		
A:Status: preliminary; translated from GB/EMBL/DBJ		
A:Molecule type: DNA		
A:Residues: 1-918 <GB>		
A:Cross-references: EMBL:AF003136; PIDN:AA93636.1; GSPDB:GN00019; CESP:F28B3.5		
A:Experimental source: strain Bristol N2; clone F28B3		
C:Genetics:		
A:Gene: CESP:F28B3.5		
A:Map position: 1		
A:Introns: 85/3; 129/3; 235/3; 418/2; 482/3; 532/3; 736/3; 829/3; 850/2		
Query Match	24.4%	Score 470; DB 2; Length 918;
Best Local Similarity	33.0%	Pred. No. 5.8e-32;
Matches 131; Conservative 75; Mismatches 143; Indels 48; Gaps 14;		
QY	1	MLSTLVHTYSMRYLPSV...SVLLGTAPTYVLAMGVMRLSAFLPAREYQALDRLYC 56
DB	533	MLSTLVAVDALRPIRCSLSTSMVPEFASCAIYIGV...SWIVPRHVAQQLDMLTK 588
QY	57	VYQSMVLFPEENTGTGVOILLYGLPK--NK-----ENIIYLANHOSYDMDIVADILAIRON 110
DB	589	SYNMLCLFVEENSGVEIYHGTNEEVNKTGRKNVMSINQSNVDIIPVLAARHG 648
QY	111	ALGH---VRVVLKRGKWLPLCYGCFPAHOGCIYVRSKAKNEKMRKLSYDAGTPMY 167
DB	649	DQGNQAFRVWVKNKSHLVMEFGYTRFQHGITYRRGEGTIGAPVLRQLKMLNESDPYW 708
QY	168	LVTFPESTRYNPBOTKVLVSAQFAOAGLAVLKHVLPRIKATHVAFDCKMYLDAIYD 227
DB	709	LLTFPESTRNSAKKHLLESNRFLEKSGRQPMQNVLCPRSGGLQALDNLST-LDAIYD 767
QY	228	VTVVYV-EGKDDG-----QRESPTWFLC--KECPRIHIDRIDK 267
DB	768	VFWMYGQMRPDLGINLTIHYNLIIYRMAERGLAPMPDFCCGSOQFOKLHLDLPIPI 827
QY	268	KDVPDEGHEHRRMLHEFEELKDKMLIEFYSPDERKRRPKGSVNSKLSIKKTLPSMLI 327
DB	828	DEVPKARLELRTWTIEFTFKERIDFV-SEKPSGSGALP-----CVPIISOFLPSLTF 880
QY	328	LS-GTFAGMLMTDAGRLVYNTWIVGTLLGLCWVTIK 363

Db 881 FSAALLAPFFSRITGR-IYLLTIASSPLL-IAMLHIR 915

RESULT 2

B96780  
hypothetical protein F9E10.13 [imported] - Arabidopsis thaliana  
C:Species: Arabidopsis thaliana (mouse-ear cress)  
C>Date: 02-Mar-2001 #sequence\_revision 02-Mar-2001 #text\_change 31-Mar-2001  
C:Accession: B96780  
R:Theologias, A.; Ecker, J.R.; Palm, C.J.; Federpiel, N.A.; Kaul, S.; White, O.; Alonso,  
Chn, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Cressy, T.H.; Dewar, K.;  
ansen, N.F.; Hughes, B.; Huizar, L.  
Nature 408, 816-820, 2000  
A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.  
C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziani,  
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.  
A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,  
ker, M.; Wu, D.; Yu, G.; Traiser, C.M.; Venter, J.C.; Davis, R.W.  
A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.  
A:Reference number: A86141; MUID:21016719  
A:Accession: B96780  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-393 <STO>  
A:Cross-references: GB:AEO05173; MID:96646762; PIDN:AAF21074.1; SSPDB:GN00141  
C:Genetics:  
A:Gene: F9E10.13  
A:Map position: 1

Query Match 20.7%; Score 398.5; DB 2; Length 393;  
Best Local Similarity 29.8%; Pred. No. 2.5e-26;  
Matches 97; Conservative 68; Mismatches 117; Indels 43; Gaps 7;

12 MRYLLPSVYLGTPTATYLANGVWRILSAFLPARFYQALDRLCYQOSMWL----- 63  
Db 23 LRGLMILIVFLSTAPMFL-----YFAP---IALGLRLSLVQOSKRVYSLIFGLM 70  
Qy 64 -----FFENYTGVOILYGLDLPKKNENIILANHOQSTVDIVADILAIROMALGHVRYV 118  
Db 71 LALMPLEFYLVNGTGVVSSDIIPVEKRVLLIANHREVDVMWMLNIALRGCGGIIYV 130  
Qy 119 LKEGLKMLPLGYCYFAOHGIGYVRSKAFENKEMRNKIQSYVDAGTPMYLVIPFGTRYN 178  
Db 131 LKSSLMKLPFGWGFHYLDEFIPVERKREVDPEVLQMLSFKDOPEPLMALFPEGIDFT 190  
Qy 179 PEQTKVLASQAFAORGLAVLKHVLPRIKATHVAFPCMKNYLDATYDVVVVEGKODG 238  
Db 191 EEKK---RSOKFAVEGLPALSNVLLPRTKGFVGLVNLNSIDAVDITATKPR-- 244  
Qy 239 GQRESPTMTFELCKECP-KIHIDRIDKKDVEPEOEHRMRWLHERFEIKDKMLIEFY 297  
Db 245 -----CSPFMDNVFGTDPSEVHHVRRYLKLEIFANEASSAMLMDSKFLDKLLSDF-- 297  
Qy 298 SPDEPRKRRFGKSVNSKLITKTL 322  
Db 298 ----NAOGKFPNQRPEELSVLKCI 318

RESULT 3

S52645  
probable 1-acyl-glycerol-3-phosphate acyltransferase - maize  
C:Species: Zea mays (maize)  
C>Date: 14-Jul-1995 #sequence\_revision 21-Jul-1995 #text\_change 19-May-2000  
C:Accession: S52645  
R:BROWN, A.P.; Coleman, J.; Tomney, A.M.; Watson, M.D.; Slabbs, A.R.  
Plant Mol. Biol. 26, 211-223, 1994  
A:Title: Isolation and characterization of a maize cDNA that complements a 1-acyl sn-gly  
her acyltransferases.  
A:Reference number: S52645; MUID:95035993  
A:Accession: S52645  
A:Status: preliminary  
A:Molecule type: mRNA

A:Residues: 1-374 <BRO>  
A:Cross-references: EMBL:Z29518; MID:9575959; PIDN:CAA82638.1; PID:9575960  
C:Superfamily: probable membrane protein YBR042c

Query Match 17.0%; Score 328; DB 2; Length 374;  
Best Local Similarity 29.8%; Pred. No. 2.7e-20;  
Matches 92; Conservative 64; Mismatches 121; Indels 32; Gaps 8;

29 VLAMGWRLLSA-----FLPAR-----FYQALDRLCYQOSWVLFPEFNYTGVQ 73  
Db 9 VLPGLFLSLGLLYNAIQAVLFVITRPFSSFTIRIRFLAELIMQLQVAVDMWAGVK 68  
Qy 74 ILXYD----LPKKNENIILANHOQSTVDIVADILAIROMALGHVRYLKEGLKMLPLY 129  
Db 69 VOLHDEFTYSMCKEHALIISNRSDIDWILGIMLAQRSCGLSTLAVMKSSKFLPYI 128  
Qy 130 G-CYFAOHGIGYVRSKAFENKEMRNKIQSYVDAGTPMYLVIPFGTRYNPEQTKVLSA 187  
Db 129 GWSMWFAYE--LFLERSWAKDEKTLKWLQRLKDFPRPFMLALFEGTRFPAR--LLA 183  
Qy 188 SOAFAOGLAVLKHVLPRIKATHVAFPCMKNYLDATYDVVVEGKODGQRESPTM 247  
Db 184 AOETAASOGLPAPKRVLLPRTKGFVSAVSIMRDEVPALYDTVLP-KDS-----POPTM 237  
Qy 248 TEFCKECPKIHIDRIDKKDVEPEOEHRMRWLHERFEIKDKMLIEFESPDERRKRF 307  
Db 238 LRLKGSSVYIHWKMRKRAMSEMPKSDSDVSKWCKDIFVANDALLDKLATGTDEEIRP 297  
Qy 308 PGKSVNSKL 316  
Db 298 IGRPVKSL 306

RESULT 4

S60478  
probable 1-acyl-sn-glycerol-3-phosphate acyltransferase - Limnanthes douglasii  
C:Species: Limnanthes douglasii  
C>Date: 19-Mar-1997 #sequence\_revision 19-Mar-1997 #text\_change 17-Mar-2000  
C:Accession: S60478  
R:BROWN, A.P.; Brough, C.L.; Kroon, J.T.M.; Slabbs, A.R.  
Plant Mol. Biol. 29, 267-278, 1995  
A:Title: Identification of a cDNA that encodes a 1-acyl-sn-glycerol-3-phosphate acylt  
A:Reference number: S60477; MUID:96046746  
A:Accession: S60478  
A:Status: nucleic acid sequence not shown; translation not shown  
A:Molecule type: mRNA  
A:Residues: 1-377 <BRO>  
A:Cross-references: EMBL:Z48730; MID:91067137; PIDN:CAA88620.1; PID:91067138  
A:Note: the nucleotide sequence was submitted to the EMBL Data Library, March 1995  
C:Superfamily: probable membrane protein YBR042c

Query Match 16.0%; Score 307; DB 2; Length 377;  
Best Local Similarity 25.5%; Pred. No. 1.7e-18;  
Matches 91; Conservative 85; Mismatches 131; Indels 50; Gaps 13;

37 LLSAFLPARFY---OALDRLCYQOSWV-----LFFENYTGVOILYGLDLPK--- 82  
Db 21 LVNFIQAVFYVLRPIKDTYRRINTIYVALLMELWVLDWAGVAVQVLTDTESRL 80  
Qy 83 -NKENIILANHOQSTVDIVADILAIROMALGHVRYLKEGLKMLPLYG--CYFAOHGI 139  
Db 81 MGEHALDICNHRSDIDWILGIMVLAQRSCGLSSIAVVKSSKFLPYIGWMSFSEY--L 138  
Qy 140 YVKSARKNEKEMRNKIQSYVDAGTPMYLVIPFGTRYNPEQTKVLSA 198  
Db 139 FLERWAKADEMTLKSGLRLNDFPPEFLALFVETRP---TRAKLLAAOEYASAPLP 194  
Qy 199 VLKHLVPRIKATHVAFPCMKNYLDATYDVVVVEGKODGQRESPTMTFELCKECPKI 258  
Db 195 VPRNVLPRTKGFVSAVSNMRSFVPALYDTVAIP-----KITEQPTMLALFRKSSV 248



```

Db 122 -GAGMAMSSGSYIFLJNRNFENDKRVLEIRIYKYSGSEKKQIILLFAGCTDGGKRAATL-- 178
Qy 187 ASQAPAAORGLAVLKHYLTPTRIKTHAFCQMK--NLIDAIYDVTYVYBEKDD3GGQRRES 244
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db 179 -SDAFADKNGLPRIEYVHPPTTFQFKFLMELMKENKIKYVDITLITVSG-----TIV 230
Qy 245 PTMTPEFLCKECP-KITHIDRIDKNDVPEQDHRMRMLHERFEIKDKMLLEFYSPD 300
Db 231 DTEKLLAGNFPDVVHLDPVKRYKRLDEIPL-NGEGCEKMLTDLAMATKRRILKPFYIOE 286

```

RESULT 8  
T079936  
Probable glycerol-3-phosphate O-acyltransferase (EC 2.3.1.15) - rape  
N:Alternate names: 1-acyl-sn-glycerol-3-phosphate acyltransferase  
C:Species: *Brassica napus* (rape)  
C:Date: 21-May-1999 #sequence\_rev1sion 21-May-1999 #text\_change 20-Jun-2000  
C:Accession: T079936  
R:Brough, C.L.  
submitted to the EMBL data library, June 1995  
A:Reference number: Z16230  
A:Accession: T079936  
A:Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-311 <PRO>  
A:Cross-references: EMBL:Z43860; NID:g1149594; PIDD:CAA90019.1  
A:Experimental source: cv. 'jet neuf'; embryo  
A:Superfamily: Probable membrane protein YBR042c  
;Keywords: acyltransferase; coenzyme A

	Query Match	12.4%	Score 239	DB 2	Length 311
	Best Local Similarity	29.8%	Pred No. 9.2e-13		
	Matches 59	Conservative 46	Mismatches 79	Indels 14	Gaps 5
QY	97	VDVIADILAIRQNALGHRVYLKSGLTKMLPLYG--CYFAHQGIGIVYKRSAKFENKEMRN	154		
		: : : : :   : : : : :   : : : : :   : : : : :   : : : : :   : : : : :			
Db	17	ISGLVYVNLQ--RSGCLSGALAVMKSSKFLPVIGMSMFSEX--LFLEERNMKDSESTLKS	73		
QY	155	KLGASVDAGTPLYLVIFPGFTGRYNEQRTKVSASQAFAQROBLAVLTKHVLTPRIATFNH	214		
		: : : : :   : : : : :   : : : : :   : : : : :   : : : : :   : : : : :			
Db	74	GIQRINDPRLPRLALFLVYGTGRTFAK--LKAADYEAASSELPAVRNVLIRTGQFYSA	130		
QY	215	FDCMKYLDAYDVYVVEKGGKDGQRRSEPMTEFLCKECPRIHIIHDIRDKVAPDEQO	274		
		: : : : :   : : : : :   : : : : :   : : : : :   : : : : :   : : : : :			
Db	131	VSNMSEFPAIDMTVAIP-----KTSPPRLMLRLFKQGRSVVYHIIHCHSMKDLFESE	184		
QY	275	EHRRLVHERFEIKDKML	292		
		: : : : :   : : : : :   : : : : :   : : : : :   : : : : :   : : : : :			
Db	185	DEIAQWCRDQFVTKDALL	202		

RESULT 9  
 T06755  
 Probable glycerol-3-phosphate O-acyltransferase (EC 2.3.1.15) - Arabidopsis thaliana  
 N:Alternate names: protein F15B8.160  
 C:Species: Arabidopsis thaliana (mouse-ear cress)  
 C:Date: 23-Apr-1999 #sequence\_revision 23-Apr-1999 #text\_change 19-May-2000  
 C:Accession: T06755  
 R:Queller, F.; Benes, V.; Rechmann, S.; Borkova, D.; Ansoerge, W.; Salanoubat, M.; Mewes  
 Submitted to the Protein Sequence Database, April 1999  
 A:Reference number: 215794  
 A:Accession: T06755  
 A:Molecule type: DNA  
 A:Residues: 1-310 <QUE>  
 A:Cross-references: EMBL:AL049660; GSPDB:GN00061; ATSP:F15B8.160  
 A:Experimental source: cultivar Columbia; BAC clone F15B8  
 C:Genetics:  
 A:Gene: ATSP:F15B8.160  
 A:Map position: 3  
 A:Introns: 26/3; 46/3; 72/3; 125/3; 167/3; 198/3; 227/3  
 C:Superfamily: probable membrane protein YBR042c  
 ;Keywords: acyltransferase; coenzyme A

Query Match	12.4%	Score 238.5;	DB 2;	Length 310;
Best Local Similarity	24.8%	Pred. No. 1e-12;		
Matches 67; Conservative	64;	Mismatches 106;	Indels 33;	Gaps 8;

QY	108	ROALGHVRYVLEKELMKLPYLG--CYAAGHGIVYKRSAKFNENKRNKIOQVVDGTP	165
		:         :   :   :   :   :   :   :   :   :   :   :   :	
Db	27	RSGCGSALAVWKKSSKFLPVIGSSMMFSEY--LFLENNAKKDSTLKSGLQRLSDPPRP	84
QY	166	MYLVIFPEGSTRYNPEQTKVLSASQAFAAORGLAVLKHVLPRIKATHVAEDCMKNYDAI	225
		:   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :	
Db	85	FVLALFVETGRTPEAK--LKAOEYASSELPIPRNVILPRTGFPVASNMRSFVPAL	145
QY	226	YDVAVVYEGKDDGGORRSPMTIELCECKEIKIHIIHIDRIKRVPEQGEIMRWLHERF	285
Db	142	YDMATYTIIP-----KTSPPPLMLRLFQGPBVHVIHKSHCMKLPESDDAIAQMCRDQF	195
QY	286	EIKDKMLIEFESPDPERRRRRFPKGSVNSKISIKITPLSMILSG---LTAGML-----	336
		:   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :	
Db	196	VAKDALLDKNIAA-----DTEPGQ--QEQNICRPIKSLAVLVSMACVLTIGAIKFLHW	246
QY	337	--MDAGRKLYVNMVYITGLLGLCLMWTYKA	364
		:   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :	
Db	247	AQLESMSKGIYISALGIITLCMOILRS	276

RESULT 10  
S45900  
probable membrane protein YBR042c - yeast (*Saccharomyces cerevisiae*)  
N.Alternate names: hypothetical protein YBR0412  
C.Species: *Saccharomyces cerevisiae*  
C.Date: 26-Aug-1994 #sequence\_revision 09-Sep-1994 #text\_change 22-Oct-1999  
C.Accession: S45900  
R.Andre, B.; Cziepluch, C.; Hein, C.; Jauniaux, J.C.; Urrestiarazu, A.; Visiers, S.  
submitted to the Protein Sequence Database, August 1994  
A.Reference number: S45893  
A.Accession: S45900  
A.Molecule type: DNA  
A.Residues: 1-397 <AND>  
A.Cross-references: EMBL:Z35911; NID:9536265; PIDD:CAAB4984.1; PID:9536266; GSPDB:GN00  
A.Experimental source: strain S288C  
C.Genetics:  
A.Gene: MIPS:YBR042c  
A.Map position: 2R  
C.Superfamily: probable membrane protein YBR042c  
C.Keywords: transmembrane protein  
F.12-37/Domain: transmembrane #status predicted <TM1>  
F.55-77/Domain: transmembrane #status predicted <TM2>  
F.134-150/Domain: transmembrane #status predicted <TM3>  
F.372-390/Domain: transmembrane #status predicted <TM4>

	Query Match	Similarity	12.08;	Score	231.5;	DB	2;	Length	397;
	Best Local	Similarity	25.99;	Pred.	No.	5.7e-12;			
	Matches	88;	Conservative	55;	Mismatches	136;	Indels	61;	Gaps
Oy	54	LVCYQASVMEFFENNYTGVQILLYGDLPEKN-----KENITLYANHOSTVDWIVADIIAI	107						
	:	: : : : : :							
	:	: : : : : :							
Dd	67	LHVAPSAVATRTTESSVPKCTFPELDLKKRILSHLSKSNSVAIICNHIQITFDWIFLMWLAY	126						
	:	: : : : : :							
Oy	108	RQNALGHRYVLKEBKLPLRLYGCTCFADHGSIYYKRSKAFNEKEMRKLL	156						
	:	: : : : : :							
Dd	127	TSLNGANVFILIKSLASIPILPGOMRNYNFIEMFSRMKAQDKITLSSLNLGLDSNARAG	186						
	:	: : : : : :							
Oy	157	-----OSYDAGT-----PMYLVIEPEGTGRYPDEOTKYLSAQARAAQ	194						
	:	: : : : : :							
Dd	187	SLACKSPERITEEGESIMNPVEIDPKQIHMYNNILPEREGINLSAD--TRQSKAK--YAAK	243						
	:	: : : : : :							
Oy	195	RGLAVLKHVLTLPRIKAHVAFCCKNKNYALDIYDVTVVYEG--KDDGGQRRESPTMTFLC	255						
	:	: : : : : :							
Dd	244	IGKKFKFNVLPHSTGTLGYSLQKLRCPSTESLYDTDIGSGVKQGEVDELTYG-LKTSFTLE	302						







---

**THIS PAGE BLANK (USPTO)**

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 28, 2002, 11:17:22 ; Search time 21.99 Seconds

(without alignments)  
640.924 Million cell updates/sec

Title: US-09-853-526-5

Perfect score: 1924

Sequence: 1 MLSTLVHTSMRYLPSVY.....YVNWIVGTLLGLMTVTKA 364

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 105224 seqs, 38719550 residues

Total number of hits satisfying chosen parameters: 105224

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt\_40:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No	Score	Query Match	Length	ID	Description
1	1871	97.2	353	1	P1CE_HUMAN
2	1568.5	81.5	354	1	P1CE_MOUSE
3	270	14.0	378	1	P1CD_RAT
4	268.5	14.0	378	1	P1CD_HUMAN
5	240	12.5	376	1	P1CC_HUMAN
6	231.5	12.0	397	1	P1B2_YEAST
7	221.5	11.5	396	1	P1B2_YEAST
8	204.5	10.6	370	1	P1Y205_HUMAN
9	182	9.5	391	1	P1Y73_CAEEL
10	164.5	8.5	310	1	P1YING_CAEEL
11	124	6.4	303	1	P1SC_YEAST
12	119	6.2	281	1	P1SC_LIMDO
13	116	6.0	266	1	P1SC_MYCPN
14	116	6.0	281	1	P1SC_LIML
15	111.5	5.8	377	1	P1VENY_FOWPY
16	111.5	5.8	488	1	P1EXON_HSV62
17	109.5	5.7	488	1	P1EXON_HSV62
18	109	5.7	240	1	P1SC_HAETN
19	104	5.4	240	1	P1SC_HELPI
20	104	5.4	308	1	P1SC_COCNU
21	102	5.3	1551	1	P1Y012_CAEEL
22	101	5.2	262	1	P1C1_CAEEL
23	99.5	5.2	237	1	P1SC_HEIRP
24	99	5.1	245	1	P1SC_SALTY
25	98	5.1	484	1	P1Y142_NPVOP
26	97	5.0	245	1	P1SC_ECOTI
27	94.5	4.9	282	1	P1C2_CAEEL
28	93	4.8	250	1	P1SC_BORBU
29	90.5	4.7	1241	1	P1TK1_SACBA
30	90	4.7	892	1	P1RAL6_SCHPO
31	88.5	4.6	446	1	P1M64_MOUSE
32	87.5	4.5	503	1	P1CPN1_PAPHA
33	86.5	4.5	268	1	P1SC_MYCCE

34	86	4.5	493	1	SYK_ARCFU
35	85.5	4.4	160	1	ALGO_PSEAE
36	85	4.4	561	1	5NRC_HUMAN
37	84.5	4.4	255	1	P1SC_NEIMB
38	84.5	4.4	383	1	P1D6_ARATH
39	84	4.4	1272	1	Y228_METJA
40	83.5	4.3	255	1	P1SC_NEIMO
41	83.5	4.3	255	1	P1SC_NEIMA
42	83.5	4.3	503	1	CPN1_HUMAN
43	83.5	4.3	2386	1	RAD3_SCHPO
44	83	4.3	318	1	YOK3_CAEEL
45	83	4.3	778	1	YF05_METJA

## ALIGNMENTS

RESULT ID	1	PLCE_HUMAN	STANDARD:	PRT:	353 AA.
AC	09NU02; 09BOG4;				
DT	01-MAR-2002 (Rel. 41, Created)				
DT	01-MAR-2002 (Rel. 41, Last sequence update)				
DT	01-MAR-2002 (Rel. 41, Last annotation update)				
DE	1-acyl-sn-glycerol-3-phosphate acyltransferase epsilon (EC 2.3.1.51)				
DE	(1-AGP acyltransferase 5) (1-AGPAT 5) (Lysophosphatidic acid acyltransferase-epsilon) (LPAAT-epsilon) (1-acylglycerol-3-phosphate O-acyltransferase 5).				
DE	AGPAT5.				
GN	AGPAT5.				
OS	Homo sapiens (Human).				
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.				
OX	NCBI_TaxID=9606;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RT	Leung D.W.;				
RT	"Cloning and expression of LPAAT-epsilon.";				
RL	Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.				
RN	[2]				
RP	SEQUENCE FROM N.A.				
RA	Cohen D., Chumakov I., Blumenfeld M., Bougueleret L.;				
RL	Patent number WO9332644, 01-JUL-1999.				
RN	[3]				
RP	SEQUENCE FROM N.A.				
RC	TISSUE=Amalgam;				
RX	MEDLINE=2115917; PubMed=11230166;				
RA	Wiemann S., Well B., Wellenreuther R., Gassenhuber J., Glassl S.,				
RA	Ansorge W., Boecher M., Bloeker H., Bauersachs S., Blum H.,				
RA	Lauber J., Duesterhoeft A., Beyer A., Koehler K., Strack N.,				
RA	Mewes H.-W., Othenmaier B., Obermaier B., Tampe J., Heubner D.,				
RA	Wambutt R., Korn B., Klein M., Poustka A.;				
RT	"Towards a catalog of human genes and proteins: sequencing and				
RT	analysis of 500 novel complete protein coding human cDNAs.";				
RL	Genome Res. 11:422-435(2001).				
RN	[4]				
RP	SEQUENCE FROM N.A.				
RC	TISSUE=Placenta;				
RA	Isogai T., Ota T., Hayashi K., Sugiyama T., Otsuki T., Suzuki Y.,				
RA	Nishikawa T., Nagai K., Sugano S., Shiratori A., Sudo H.,				
RA	Wagatsuna M., Hosokiri T., Raku Y., Kodaira H., Kondo H., Sugawara M.,				
RA	Takahashi M., Chiba Y., Ishida S., Murakawa K., Ono Y., Takiguchi S.,				
RA	Watanabe S., Kimura K., Murakami K., Ishii S., Kawai Y., Saito K.,				
RA	Yamanoto J., Wakatsuki A., Nakamura Y., Nagahara K., Masuno Y.,				
RA	Nimomiya K., Iwayanagi T.;				
RT	"NEO human cDNA sequencing project.";				
RL	Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.				
CC	- FUNCTION: CONVERTS LYOPHOSPHATIDIC ACID (LPA) INTO PHOSPHATIDIC				
CC	ACID BY INCORPORATING ACYL MOIETY AT THE 2 POSITION (BY				
CC	SIMILARITY).				
CC	- CATALYTIC ACTIVITY: Acyl-CoA + 1-acyl-sn-glycerol 3-phosphate =				
CC	CoA + 1,2-diacyl-sn-glycerol 3-phosphate.				
CC	- PATHWAY: SECOND STEP IN DE NOVO PHOSPHOLIPID BIOSYNTHESIS.				
CC	- SUBCELLULAR LOCATION: Integral membrane protein (potential).				

CC -1- SIMILARITY: BELONGS TO THE 1-ACYL-SN-GLYCEROL-3-PHOSPHATE  
 CC ACYLTRANSFERASE FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL: AF375789; AKS4809.1; ALT\_INT.  
 DR EMBL: AL136587; CAB6622.1; ALT\_INT.  
 DR EMBL: AK002072; BAA92069.1; -  
 DR InterPro: IPR002123; Acyltransferase.  
 DR Pfam: PF01553; Acyltransferase; 1.  
 KM Phospholipid biosynthesis; Transferase; Acyltransferase;  
 KM Transmembrane.  
 FT TRANSMEM 7 29 POTENTIAL.  
 FT TRANSMEM 44 66 POTENTIAL.  
 FT TRANSMEM 334 351 POTENTIAL.  
 FT CONFLICT 145 145 L -> V (IN REF. 2).  
 SQ SEQUENCE 353 AA; 40813 MW; A05B1FA246CE1B64 CRC64;

Query Match 97.2%; Score 1871; DB 1; Length 353;  
 Best Local Similarity 100.0%; Pred. No. 1.3e-151;  
 Matches 353; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 12 MRYLLPSVLLGTAPTYVLAWGVRLLSAFLPARFYQALDRLCYQSVLFFENYTG 71  
 DB 1 MRYLLPSVLLGTAPTYVLAWGVRLLSAFLPARFYQALDRLCYQSVLFFENYTG 60  
 QY 72 VQILLYGDLPKKNENITYLANHSTVDIMYADILAIKQNALGHRVYLKESGLMPLXGC 131  
 DB 61 VQILLYGDLPKKNENITYLANHSTVDIMYADILAIKQNALGHRVYLKESGLMPLXGC 120  
 QY 132 YFAOHGCIYVRSKAKFENKEKRNKLSQSYVDAGTPMYLVIEPEGTRVNPEDTKVLSAQAF 191  
 DB 121 YFAOHGCIYVRSKAKFENKEKRNKLSQSYVDAGTPMYLVIEPEGTRVNPEDTKVLSAQAF 180  
 QY 192 AAQAGLAVLKHVLPRIKATKHAFAFDCKKNYLDATYDVTYVEGKDDGQRRSPPTMTEFL 251  
 DB 181 AAQAGLAVLKHVLPRIKATKHAFAFDCKKNYLDATYDVTYVEGKDDGQRRSPPTMTEFL 240  
 QY 252 CKECPKIHIDRIKDKDVEEODEHMRRLHEREFEIKDKMLIEFYESPDPERRKRPFGKS 311  
 DB 241 CKECPKIHIDRIKDKDVEEODEHMRRLHEREFEIKDKMLIEFYESPDPERRKRPFGKS 300  
 QY 312 VNSKLSIKKTLPSMLLSGLTAGMLMTDAGRKLYVNTWYIGTLGLCAWTTIKA 364  
 DB 301 VNSKLSIKKTLPSMLLSGLTAGMLMTDAGRKLYVNTWYIGTLGLCAWTTIKA 353

RESULT 2  
 PLOC\_MOUSE  
 ID PLOC\_MOUSE STANDARD; PRT; 354 AA.  
 AC 09D1E8;  
 DT 01-MAR-2002 (Rel. 41, Created)  
 DT 01-MAR-2002 (Rel. 41, Last sequence update)  
 DE 1-acyl-sn-glycerol-3-phosphate acyltransferase epsilon (EC 2.3.1.51)  
 DE acyltransferase-epsilon 5 (1-AGPAT 5) (lysophosphatidic acid  
 DE O-acyltransferase 5).  
 GN AGPAT5.  
 OS Mus musculus (mouse).  
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=C57BL/6J; TISSUE=Embryo;

RX MEDLINE=21085660; PubMed=11217851;  
 RA Kawai J., Shingawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,  
 RA Arakawa T., Hara A., Fukunishi Y., Kono H., Adachi J., Fukuda S.,  
 RA Atzawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamataka I.,  
 RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,  
 RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,  
 RA Kleihsman W., Gaasterland T., Glasi C., King B., Kochia H.,  
 RA Kuehl P., Lewis S., Matsuo Y., Nikaide I., Pesole G., Quackenbush J.,  
 RA Schirral L.M., Staubli R., Suzuki R., Tomita M., Wagner L., Washio T.,  
 RA Sakai K., Okido T., Furuno M., Aono H., Balderelli R., Barsch G.,  
 RA Blake J., Boffelli D., Bojunga N., Carrinci P., de Bona M.F.,  
 RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,  
 RA Gustinich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,  
 RA Lyons P., Marchionni L., Mashima Y., Mazzarelli J., Mombaerts P.,  
 RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,  
 RA Sasaki H., Sato K., Schoenbach C., Seva T., Shibata Y., Storch K.-F.,  
 RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whitaker C., Wilmink L.,  
 RA Wyshak-Boris A., Yoshida K., Hasegawa Y., Kawai H., Kohsaki S.,  
 RA Hayashizaki Y.,  
 RT "Functional annotation of a full-length mouse cDNA collection."  
 RL Nature 409:685-690(2001).  
 CC -1- FUNCTION: CONVERTS LYSOPHOSPHATIDIC ACID (LPA) INTO PHOSPHATIDIC  
 CC ACID BY INCORPORATING ACYL MOIETY AT THE 2 POSITION (BY  
 CC SIMILARITY).  
 CC -1- CATALYTIC ACTIVITY: Acyl-CoA + 1-acyl-sn-glycerol 3-phosphate -  
 CC CoA + 1,2-diacyl-sn-glycerol 3-phosphate.  
 CC -1- PATHWAY: SECOND STEP IN DE NOVO PHOSPHOLIPID BIOSYNTHESIS.  
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein (Potential).  
 CC -1- SIMILARITY: BELONGS TO THE 1-ACYL-SN-GLYCEROL-3-PHOSPHATE  
 CC ACYLTRANSFERASE FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL: AK003649; BAB2915.1; -  
 DR MGD: MGI:1915880; Agpat5.  
 DR InterPro: IPR002123; Acyltransferase.  
 DR Pfam: PF01553; Acyltransferase; 1.  
 KM Phospholipid biosynthesis; Transferase; Acyltransferase;  
 KM Transmembrane.  
 FT TRANSMEM 7 29 POTENTIAL.  
 FT TRANSMEM 44 66 POTENTIAL.  
 FT TRANSMEM 335 352 POTENTIAL.  
 SQ SEQUENCE 354 AA; 40943 MW; 447BEF9249B1E800 CRC64;

Query Match 81.5%; Score 1568.5; DB 1; Length 354;  
 Best Local Similarity 81.4%; Pred. No. 6.5e-126;  
 Matches 288; Conservative 39; Mismatches 26; Indels 1; Gaps 1;

QY 12 MRYLLPSVLLGTAPTYVLAWGVRLLSAFLPARFYQALDRLCYQSVLFFENYTG 71  
 DB 1 MRYLLPSVLLGTAPTYVLAWGVRLLSAFLPARFYQALDRLCYQSVLFFENYTG 60  
 QY 72 VQILLYGDLPKKNENITYLANHSTVDIMYADILAIKQNALGHRVYLKESGLMPLXGC 131  
 DB 61 VQILLYGDLPKKNENITYLANHSTVDIMYADILAIKQNALGHRVYLKESGLMPLXGC 120  
 QY 132 YFAOHGCIYVRSKAKFENKEKRNKLSQSYVDAGTPMYLVIEPEGTRVNPEDTKVLSAQAF 191  
 DB 121 YFAOHGCIYVRSKAKFENKEKRNKLSQSYVDAGTPMYLVIEPEGTRVNPEDTKVLSAQAF 180  
 QY 192 AAQAGLAVLKHVLPRIKATKHAFAFDCKKNYLDATYDVTYVEGKDDG-QGRRSPPTMTEFL 250  
 DB 181 AAQAGLAVLKHVLPRIKATKHAFAFDCKKNYLDATYDVTYVEGKDDG-QGRRSPPTMTEFL 240  
 QY 251 CKECPKIHIDRIKDKDVEEODEHMRRLHEREFEIKDKMLIEFYESPDPERRKRPFGK 310  
 DB 241 CKECPKIHIDRIKDKDVEEODEHMRRLHEREFEIKDKMLIEFYESPDPERRKRPFGK 300



Db 133 GMMWTFTEM--VFQSRKMEQDQKRVATSDHMLRDYPERKQYFFLLCBSTRTEKKHEI--- 187

QY 188 SOAFQAQGLAVLKVLTLPRTKATHVAFDCMKNTLDAIYDVTYVYEGSKDQDGRRESPTM 247

Db 188 SMOVARAAGRLRLKIHLLPRTKGAIVTRSLRVNVSAVYQCTLMF-----ENNENPTL 240

QY 248 TEFLICEKSRPHIHILRIIDKKDVEEEOBHMRMLHEFEFLKDKMLIEFY-----ESP-D 300

Db 241 LGVLNGKKYHNDLVYRRPLIEDIEDDEDCSAMLHKLYOEKDAFOEYELVTGTFPELPMV 300

QY 301 PERKKRRFGKSYN 313

Db 301 PPRR---PWTLLVN 310







DR EMBL: AE000461; AAC76860.1; -.  
 DR PIR: S40808; S40808.  
 DR Ecogene: EG11833; y1ng.  
 DR InterPro: IPR002123; Acyltransferase.  
 DR Pfam: PF01553; Acyltransferase; 1.  
 KW Hypothetical protein; Complete proteome.  
 SQ SEQUENCE 310 AA; 36289 MW; 9F8E3F52EB0B186E CRC64;

Query Match 8.5%; Score 164.5; DB 1; Length 310;  
 Best Local Similarity 20.5%; Pred. No. 1.1e-06;  
 Matches 66; Conservative 69; Mismatches 128; Indels 59; Gaps 13;

QY 1 MLISVLHYHSMRYLLPSVVLIGTAPTYVLAWGVLLSAPFLPARYQALDRLCYVYQS 60  
 DB 19 LILSTVLTLTVTTCGVPPIITAGIKVLLPVPYIMKVSF-----CDFMYTCWCEG 70  
 QY 61 MWLFFENYTGVOILLYGDLPRKNEN-IYLANHSTVDWIADILAIKONALGHV---R 116  
 DB 71 LAVLLHLN-PHLQWEVHGLEJLSKKNWYLLICNHRSMADIVLCVLFK-----H1PMK 124  
 QY 117 YVLEKELKPLTYG--C-----YFAQHGITYKRSKAKNEKEMRNKLOSVDAGT-PMY 167  
 DB 125 YELKQOLAVVPELGACWSLMDPEFMKRYRAYLLRHPERKGVETTRSCSEKFRLLPPT 184  
 QY 168 LVIFPEGTYRNPEDQFVLSASQAFNAORGLAVLKHLVPRIKATHVAFQOMKNYLDAIYD 227  
 DB 185 IYNFVBSGFSFTQEK-----HQOTHSTFQNLPLPKAGIAMALNVLGKQPFKLK 233  
 QY 228 VTWVEYEGKDDGQRRSPMTETFLCECPKIHIDRIDKDVPE-----QEH 276  
 DB 234 VILCY--PDNNRQ-----PEFDMLSGKLTIRIVH---VDLPPLADLHGDIYDKSFKRH 283  
 QY 277 MRRWLHERREIDKMLIEFYES 298  
 DB 284 FOQWLMSLWQEKDRLLTSLMS 305

## RESULT 11

PLSC\_YEAST STANDARD; PRT; 303 AA.  
 ID PLSC\_YEAST STANDARD; PRT; 303 AA.  
 AC P33333;  
 DT 01-FEB-1994 (Rel. 28; Created)  
 DT 01-FEB-1994 (Rel. 28; Last sequence update)  
 DT 15-DEC-1998 (Rel. 37; Last annotation update)  
 DE Probable 1-acyl-sn-glycerol-3-phosphate acyltransferase (EC 2.3.1.51)  
 DE (1-AGP acyltransferase) (LPAAT).  
 DE acyltransferase) (LPAAT).  
 GN SLG1 OR YDL052C.  
 OS Saccharomyces cerevisiae (Baker's Yeast).  
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;  
 OC Saccharomycetales; Saccharomycetaceae; Saccharomycos.  
 OX NCB1\_TaxID=4932;  
 RN NCB1\_TaxID=4932;  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=94012814; PubMed=8408076;  
 RA Nagiec M.M., Wells G.B., Lester R.L., Dickson R.C.;  
 RT "A suppressor gene that enables Saccharomyces cerevisiae to grow  
 without making sphingolipids encodes a protein that resembles an  
 Escherichia coli fatty acyltransferase."  
 RL J. Biol. Chem. 268:22156-22163(1993).  
 [2]  
 RN SEQUENCE FROM N.A.  
 RA Urrestarazu L.A., Andre B., Vissers S.;  
 RL Submitted (JUL-1996) to the EMBL/GenBank/DBJ databases.  
 RN [3]  
 RP SEQUENCE OF 182-303 FROM N.A.  
 RA Bloeker H., Brandt P.;  
 RL Submitted (JUL-1996) to the EMBL/GenBank/DBJ databases.  
 CC -1- FUNCTION: MAY BE AN ACYLTRANSFERASE WITH AN ALTERED SUBSTRATE  
 CC SPECIFICITY THAT ENABLES IT TO USE A C-26 COA IN PLACE OF THE  
 CC C-16 OR C-18 COAS USED BY THE WILD TYPE PROTEIN.  
 CC -1- CATALYTIC ACTIVITY: Acyl-CoA + 1-acyl-sn-glycerol 3-phosphate =

CC COA + 1,2-diacyl-sn-glycerol 3-phosphate.  
 CC -1- PATHWAY: FIRST STEP IN DE NOVO PHOSPHOLIPID BIOSYNTHESIS.  
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein (Potential).  
 CC -1- SIMILARITY: BELONGS TO THE 1-ACYL-SN-GLYCEROL-3-PHOSPHATE  
 CC ACYLTRANSFERASE FAMILY.

CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC EMBL: L13282; AAA16514.1; -.  
 DR EMBL: Z74100; CAA98614.1; -.  
 DR PIR: A48600; A48600.  
 DR SGD: S0002210; SLG1.  
 DR InterPro: IPR002123; Acyltransferase.  
 DR Pfam: PF01553; Acyltransferase; 1.  
 KW Phospholipid biosynthesis; Transferase; Acyltransferase;  
 KW Transmembrane.  
 FT TRANSMEM 16 30 POTENTIAL.  
 FT VARIANT 44 44 Q -> L (IN SUPPRESSOR ALLELE SLG1-1).  
 SQ SEQUENCE 303 AA; 33887 MW; 36CBBC2659655EB CRC64;

Query Match 6.4%; Score 124; DB 1; Length 303;  
 Best Local Similarity 27.6%; Pred. No. 0.003;  
 Matches 55; Conservative 28; Mismatches 78; Indels 38; Gaps 8;

QY 10 YSMRYLLPSVVLIGTAPTYVLAWGVRL-----LSAPLPAREFQALDRLCYVYQSMYLE 64  
 DB 9 YIRSVLVVIALAGCGFYGVYASITCTLGKHLAQWTR-----CYHWKLM 58  
 QY 65 FFENYTGVOILLYGDLPRKNENIYLANHSTVDWIADILAIKONALGHV---RYVL 119  
 DB 59 L-----GLDYKVVGEENLAKKPYIMIANHGSTLDIFM-----LGRIFPGCTVNA 103  
 QY 120 KEGKMLPLTYGCFPAQHGITYKRSKAKNEKEMRNKLOSVDAGTMYLVIFEGRR-TN 178  
 DB 104 KSLKLYVPELGMFALSGYFIDRSKROEALDTLNKLENVK-NRRALWVPEGTRSYT 162  
 QY 179 PEQFKVLSAQAF-AAORG 196  
 DB 163 SELTMLPFKKGAHPHQAQG 181

## RESULT 12

PLSC\_LIMDO STANDARD; PRT; 281 AA.  
 ID PLSC\_LIMDO STANDARD; PRT; 281 AA.  
 AC Q42870; Q40120;  
 DT 15-DEC-1998 (Rel. 37; Created)  
 DT 15-DEC-1998 (Rel. 37; Last sequence update)  
 DT 15-DEC-1998 (Rel. 37; Last annotation update)  
 DE 1-acyl-sn-glycerol-3-phosphate acyltransferase (EC 2.3.1.51) (1-AGP  
 DE acyltransferase) (LPAAT).  
 DE (LPAAT).  
 GN PLSC.  
 OS Limnanthes douglasii (Douglas's meadowfoam).  
 OS Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
 OC eurosids II; Brassicales; Limnanthaceae; Limnanthes.  
 OX NCB1\_TaxID=28973;  
 RN NCB1\_TaxID=28973;  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=96028122; PubMed=7588719;  
 RA Hanke C., Wolter F.P., Coleman J., Peterek G., Frenzen M.;  
 RT "A plant acyltransferase involved in triacylglycerol biosynthesis  
 RT complements an Escherichia coli sn-1-acylglycerol-3-phosphate  
 RL acyltransferase mutant."  
 RL Eur. J. Biochem. 232:806-810(1995).  
 [2]





---

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 28, 2002, 11:16:55 ; Search time 65.95 Seconds

(without alignments)  
954.817 Million cell updates/sec

Title: US-09-853-526-5  
Perfect score: 1924  
Sequence: 1 MLSTLVHTYSMTYSLPSVY.....YVNTWYGTLLGLMTYTIKA 364

Scoring table: BIOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 562222 seqs, 172994929 residues

Total number of hits satisfying chosen parameters: 562222

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

SPTREMBL\_19:\*  
1: sp\_archaea:\*  
2: sp\_bacteria:\*  
3: sp\_fungi:\*  
4: sp\_human:\*  
5: sp\_invertebrate:\*  
6: sp\_mammal:\*  
7: sp\_mhc:\*  
8: sp\_organelle:\*  
9: sp\_phage:\*  
10: sp\_plant:\*  
11: sp\_rodent:\*  
12: sp\_virus:\*  
13: sp\_vertebrate:\*  
14: sp\_unclassified:\*  
15: sp\_virus:\*  
16: sp\_bacteriophage:\*  
17: sp\_archaeal:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1924	100.0	364	4 Q9BQ4	Q9BQ4 homo sapien
2	1868	97.1	353	4 Q9NUQ2	Q9NUQ2 homo sapien
3	1568.5	81.5	354	11 Q9D1E8	Q9D1E8 mus musculu
4	470	24.4	386	5 Q9SR12	Q9SR12 caenorhabdi
5	398.5	20.7	373	10 Q9SSH0	Q9SSH0 arabidopsis
6	398.5	20.7	373	10 Q9C9P8	Q9C9P8 arabidopsis
7	328	20.3	375	10 Q9LHNA	Q9LHNA arabidopsis
8	309.5	16.1	374	10 Q41745	Q41745 zea mays (m
9	307	16.0	390	10 Q9XFW4	Q9XFW4 brassica na
10	305.5	15.9	377	10 Q40119	Q40119 limnanthes
11	305.5	15.9	306	10 Q9SDN3	Q9SDN3 prunus dulc
12	272.5	14.2	386	5 Q9VVS1	Q9VVS1 drosophila
13	270	14.0	378	11 Q924S1	Q924S1 rattus norv
14	267	13.9	380	5 Q9VVA9	Q9VVA9 drosophila
15	263	13.7	376	10 Q9SYC8	Q9SYC8 arabidopsis
16	251	13.0	316	11 Q9DB84	Q9DB84 mus musculu

17	249.5	13.0	350	3 Q94361	Q94361 schizosacch
18	242.5	12.6	344	5 Q16526	Q16526 caenorhabdi
19	239	12.4	311	10 Q39317	Q39317 brassica na
20	238.5	12.4	310	10 Q9SVX9	Q9SVX9 arabidopsis
21	223	11.6	314	11 Q9D517	Q9D517 mus musculu
22	218.5	11.4	399	5 Q23087	Q23087 caenorhabdi
23	212	11.0	439	5 Q20800	Q20800 caenorhabdi
24	205.5	10.7	370	11 Q91YX5	Q91YX5 mus musculu
25	189.5	9.8	426	5 Q01882	Q01882 caenorhabdi
26	189.5	9.8	428	5 Q95R03	Q95R03 caenorhabdi
27	184	9.6	295	16 Q91657	Q91657 pseudomonas
28	153	8.0	304	10 Q915A6	Q915A6 pseudomonas
29	145	7.5	359	10 Q9FF57	Q9FF57 arabidopsis
30	131.5	6.8	363	5 Q19221	Q19221 caenorhabdi
31	130	6.8	294	16 Q9KVV5	Q9KVV5 vibrio chol
32	124	6.4	248	2 Q30849	Q30849 salmonella
33	116	6.0	281	10 Q9SD02	Q9SD02 limnanthes
34	114.5	6.0	247	16 Q9X219	Q9X219 thermotoga
35	113.5	5.9	385	5 Q9NFP1	Q9NFP1 leishmania
36	113	5.9	262	16 Q98G17	Q98G17 rhizobium 1
37	112	5.8	241	16 Q9CPE2	Q9CPE2 pasteurella
38	112	5.8	258	16 Q9HW50	Q9HW50 pseudomonas
39	109.5	5.7	235	2 Q9KX14	Q9KX14 rhodobacter
40	109.5	5.7	488	12 Q9WT03	Q9WT03 human herpes
41	108	5.6	261	16 Q9KPE3	Q9KPE3 vibrio chol
42	105.5	5.5	249	2 Q9EY25	Q9EY25 xanthomonas
43	105	5.5	260	16 Q92MG4	Q92MG4 rhizobium m
44	104	5.4	344	10 Q9L1Y4	Q9L1Y4 brassica na
45	102.5	5.3	291	3 Q917C5	Q917C5 emeticella

#### ALIGNMENTS

RESULT 1  
ID Q9BQ4 PRELIMINARY: PRT: 364 AA.  
AC Q9BQ4;  
DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)  
DT 01-JUN-2001 (TREMBLrel. 19, Last annotation update)  
DE HYPOTHETICAL 42.1 KDA PROTEIN (LYSOPHOSPHATIDIC ACID  
DE ACYLTRANSFERASE-EPSILON) (EC 2.3.1.51).  
GN DKFZP761C222 OR LPAAT-E.  
OS Homo sapiens (Human)  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
OX NCBI\_TaxID=9606;  
RN (1)  
RP SEQUENCE FROM N.A.  
RC TISSUE=AMGDALAE;  
RX MEDLINE=21154917; PubMed=11230166;  
RA Wiemann S., Well B., Wellenreuther R., Gassenhuber J., Glassl S.,  
RA Ansgore W., Boecker M., Bloeker H., Bauersachs S., Blum H.,  
RA Lauber J., Duesterhoeft A., Beyer A., Koehrer K., Strack N.,  
RA Mewes H.W., Ottensmeyer F., Obermaier B., Tampe J., Heubner D.,  
RA Wambutt R., Korn B., Klein M., Poustka A.,  
RA "Towards a Catalog of Human Genes and Proteins: Sequencing and  
RT Analysis of 500 Novel Complete Protein Coding Human cDNAs";  
RN genome Res. 11:422-435(2001).  
RN (2)  
RP SEQUENCE FROM N.A.  
RA Leung D.W.;  
RT "Cloning and expression of LPAAT-epsilon";  
RT Submitted (May-2001) to the EMBL/Genbank/DBJ databases.  
DR EMBL: AL136587; CAB66522.1; -;  
DR EMBL: AF375789; AAK54809.1; -;  
DR Interpro: IPR002123; Acyltransferase.  
DR Pfam: PF01553; Acyltransferase; 1.  
KW Hypothetical protein; transferase; Acyltransferase.  
SQ SEQUENCE 364 AA: 42072 MW: 90A0F87FC7C8081 CRC64;

Query Match 100.0%; Score 1924; DB 4; Length 354;  
Best Local Similarity 100.0%; Pred. No. 1.8e-167;  
Matches 364; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MLSTVHTSMRYLLPSVLLGAPTYVLAKGWRLLSAFLPARFYQALDDRLCYGQS 60  
DB 1 MLSTVHTSMRYLLPSVLLGAPTYVLAKGWRLLSAFLPARFYQALDDRLCYGQS 60  
QY 61 MVLFFENYTVQVILLTGDPKPKNKENTIIYLANHSTYDWMIVADILAIRQNALGHVRYLK 120  
DB 61 MVLFFENYTVQVILLTGDPKPKNKENTIIYLANHSTYDWMIVADILAIRQNALGHVRYLK 120  
QY 121 EGIKMLPLCYCTFAQHGIGYVKKSAKNEKEMRNKLOSVDAGTPMYLVFPESGRNPE 180  
DB 121 EGIKMLPLCYCTFAQHGIGYVKKSAKNEKEMRNKLOSVDAGTPMYLVFPESGRNPE 180  
QY 181 QTVLASQAFAAORGAVLKHVLTPIKATHVAFDCMKNYLDAIYDVTYVEGKDDGGQ 240  
DB 181 QTVLASQAFAAORGAVLKHVLTPIKATHVAFDCMKNYLDAIYDVTYVEGKDDGGQ 240  
QY 241 RRESPMTTEFLKCECPKIHIDRIDKDVPEOEHRMRLHEFEIKDKMLIEFESPD 300  
DB 241 RRESPMTTEFLKCECPKIHIDRIDKDVPEOEHRMRLHEFEIKDKMLIEFESPD 300  
QY 301 PERKKRPKGSVNSKLSIKKTLPSMLISGLTGMMLTMDAGRKLYVTWITYGTLGLWV 360  
DB 301 PERKKRPKGSVNSKLSIKKTLPSMLISGLTGMMLTMDAGRKLYVTWITYGTLGLWV 360  
QY 361 TIRKA 364  
DB 361 TIRKA 364

RESULT 2  
Q9NU02 PRELIMINARY; PRT: 353 AA.

AC Q9NU02; 01-OCT-2000 (Tremblrel. 15, Created)  
DT 01-OCT-2000 (Tremblrel. 15, Last sequence update)  
DE CDNA FLJ11210 FIS, CLONE PLACE1007954.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=PLACENTA;  
RA Isogai T., Ota T., Hayashi K., Sugiyama T., Otsuki T., Suzuki Y.,  
Nishikawa T., Nagai K., Sugano S., Shiratori A., Sudo H.,  
Wagatsuna M., Hosokiri T., Kaku Y., Kodaira R., Kondo H., Sugawara M.,  
Takatsuchi M., Chida Y., Ishida S., Murakawa K., Ono Y., Takiguchi S.,  
Watanabe S., Kimura K., Murakami K., Ishii S., Kawai Y., Saito K.,  
Yamamoto J., Wakamatsu A., Nakamura Y., Nagahara K., Masuno Y.,  
RA Minomiya K., Iwayanagi T.,  
RT "NEO human cDNA sequencing project."  
RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.  
DR EMBL: AK002072; BAA92069.1; -  
DR InterPro: IPR002123; Acyltransferase.  
DR Pfam: PF01553; Acyltransferase; 1  
SQ SEQUENCE 353 AA; 40799 MW; 0A8A1A42361F1B14 CRC64;

Query Match 97.1%; Score 1868; DB 4; Length 353;  
Best Local Similarity 99.7%; Pred. No. 2.3e-162;  
Matches 352; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 12 MRYLLPSVVLGTAPTYVLAKGWRLLSAFLPARFYQALDDRLCYCYOQNVLEFFENYTG 71  
DB 1 MRYLLPSVVLGTAPTYVLAKGWRLLSAFLPARFYQALDDRLCYCYOQNVLEFFENYTG 71  
QY 72 VQILLTGDPKPKNKENTIIYLANHSTYDWMIVADILAIRQNALGHVRYLKEGKMLPLPYGC 131  
DB 72 VQILLTGDPKPKNKENTIIYLANHSTYDWMIVADILAIRQNALGHVRYLKEGKMLPLPYGC 131

DB 61 VQILLTGDPKPKNKENTIIYLANHSTYDWMIVADILAIRQNALGHVRYLKEGKMLPLPYGC 120  
QY 132 YFAQHGIGYVKKSAKNEKEMRNKLOSVDAGTPMYLVFPESGRNPEQTKVLSAQA 191  
DB 121 YFAQHGIGYVKKSAKNEKEMRNKLOSVDAGTPMYLVFPESGRNPEQTKVLSAQA 180  
QY 192 AAORGAVLKHVLTPIKATHVAFDCMKNYLDAIYDVTYVEGKDDGGQRRSPMTTEFL 251  
DB 181 AAORGAVLKHVLTPIKATHVAFDCMKNYLDAIYDVTYVEGKDDGGQRRSPMTTEFL 240  
QY 252 CKECPRKIHIDRIDKDVPEOEHRMRLHEFEIKDKMLIEFESPDPERKKRPKGS 311  
DB 241 CKECPRKIHIDRIDKDVPEOEHRMRLHEFEIKDKMLIEFESPDPERKKRPKGS 300  
QY 312 VNSKLSIKKTLPSMLISGLTGMMLTMDAGRKLYVTWITYGTLGLWYTIKA 364  
DB 301 VNSKLSIKKTLPSMLISGLTGMMLTMDAGRKLYVTWITYGTLGLWYTIKA 353

RESULT 3  
Q9DIE8 PRELIMINARY; PRT: 354 AA.

AC Q9DIE8; 01-JUN-2001 (Tremblrel. 17, Created)  
DT 01-JUN-2001 (Tremblrel. 17, Last sequence update)  
DE 1110013A05RIK PROTEIN.  
GN 1110013A05RIK.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCBI\_Taxid=10090;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=C57BL/6J; TISSUE=EMBRYO;  
RX MEDLINE=21085660; PubMed=11217651;  
RA Kawai J., Shinagawa A., Shidara K., Yoshino M., Itoh M., Ishii Y.,  
Aizawa K., Hara A., Fukunishi Y., Kono H., Adachi J., Fukuda S.,  
Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,  
Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,  
RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiya H.,  
Kuehl P., Lewis S., Matsuo Y., Nikaide I., Pesole G., Quackenbush J.,  
Schröter L.M., Staudt F., Suzuki R., Tomita M., Wagner L., Washio T.,  
RA Sakai K., Okido T., Furuno M., Aono H., Balderelli R., Barsh G.,  
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,  
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamija M., Lee N.H.,  
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,  
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,  
RA Sasaki H., Sato K., Schenbach C., Seya T., Shibata Y., Storch K.-F.,  
RA Suzuki H., Toyooka K., Wang K.H., Weitz K., Whitaker C., Williams L.,  
RA Wysshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohlsuk S.,  
RA Hayashizaki Y.,  
RT "Functional annotation of a full-length mouse cDNA collection."  
RL Nature 409:685-690(2001).  
DR EMBL: AK003649; BAB22915.1; -  
DR MGD: MGI:1915880; 1110013A05RIK.  
DR InterPro: IPR002123; Acyltransferase.  
DR Pfam: PF01553; Acyltransferase; 1  
SQ SEQUENCE 354 AA; 40943 MW; 447EE924B91E800 CRC64;

Query Match 81.5%; Score 1568.5; DB 11; Length 354;  
Best Local Similarity 81.4%; Pred. No. 5.3e-135;  
Matches 288; Conservative 39; Mismatches 26; Indels 1; Gaps 1;

QY 12 MRYLLPSVVLGTAPTYVLAKGWRLLSAFLPARFYQALDDRLCYCYOQNVLEFFENYTG 71  
DB 1 MRYLLPSVVLGTAPTYVLAKGWRLLSAFLPARFYQALDDRLCYCYOQNVLEFFENYTG 60  
QY 72 VQILLTGDPKPKNKENTIIYLANHSTYDWMIVADILAIRQNALGHVRYLKEGKMLPLPYGC 131  
DB 72 VQILLTGDPKPKNKENTIIYLANHSTYDWMIVADILAIRQNALGHVRYLKEGKMLPLPYGC 131



Db 61 VQILLYGDLPRKNENYIYLANHSTVDWIVADMIAARODALGHVRVYLKDKLWMLPLYGF 120  
QY 132 YFAOHGIYKRSKAFNEKEMRNKLOSVDAGTPMYLIFPEEGTRVPEQTKVLSAQAF 191  
Db 121 YFQOHGIYKRSKAFNEKEMRNKLOSVDAGTPMYLIFPEEGTRVPEQTKVLSAQAF 180  
QY 192 AAGRGALVLAHVLTLPRIKATHVAFDCMKNYLDAIYDVTVVYEGKRDG-GORRESPTWTF 250  
Db 181 AAGRGALVLAHVLTLPRIKATHVAFDSKSLDAIYDVTVVYEGNEKSGKYNPSMTTF 240  
QY 251 LKCECKRIHIDRIDKDYDEOEHRMRWLHERFEIKDKMLIEFESDPERRKRPFGK 310  
Db 241 LKQCKRLHIFPRIDIRNEVEPEOEHRMRWLHERFEIKDKMLIEFYDSDPPERENKFPKG 300  
QY 311 SVNSKTSIKKTLPSMLISGLTGMMLTDAGRLVNTWYTGTLGCLWTTIKA 364  
Db 301 SVNSRSLSVKTLPSVLLGSLTAVMLMTESGRKLMGTWLYGTLGCLWTFVKA 354

## RESULT 4

Q95R12 PRELIMINARY: PRT: 386 AA.  
AC Q95R12: 01-DEC-2001 (TrEMBLrel. 19, Created)  
DT 01-DEC-2001 (TrEMBLrel. 19, last sequence update)  
DT 01-DEC-2001 (TrEMBLrel. 19, last annotation update)  
DE HYPOHETICAL 44.0 KDA PROTEIN.  
GN F28B3.9.  
OS Caenorhabditis elegans.  
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditioidea;  
OC Rhabditidae; Peloderinae; Caenorhabditis.  
OX NCBI\_TaxID=6239;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=BRISTOL N2;  
RX MEDLINE=99069613; PubMed=9851916;  
RA None;  
RT "Genome sequence of the nematode C. elegans: a platform for  
investigating biology. The C. elegans Sequencing Consortium.";  
RL Science 283:2012-2018(1998).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN=BRISTOL N2;  
RX Gelsel C., Kramer J., Smith A.;  
RA "The sequence of C. elegans cosmid F28B3.";  
RL Submitted (May-1997) to the EMBL/GenBank/DBJ databases.  
RN [3]  
RP SEQUENCE FROM N.A.  
RC STRAIN=BRISTOL N2;  
RX Waterston R.;  
RA "Direct Submission";  
RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.  
DR EMBL: AF003136; AAK93853.1; -  
KW Hypothetical protein.  
SQ SEQUENCE 386 AA: 43978 MW: AIE3020C201452A6 CRC64;

Query Match 24.4%; Score 470; DB 5; Length 386;  
Best Local Similarity 33.0%; Pred. No. 1,4e-34;  
Matches 131; Conservative 75; Mismatches 143; Indels 48; Gaps 14;

QY 1 MLISLVLHYTSMRYLPR---SVLLIGTAPTYVLAMGVWRLLSAFLPARFYQALDDRLTYC 56  
Db 1 MLISLTIANDALRPILIPCSILSLSMVPRFASCAIVIGV---SWIVPRHYAQQLDNMLYK 56  
QY 57 VYOSMVLFFENTYGVQIILYGDLPK--NK---ENIYILANHSTVDWIVADILAIKRON 110  
Db 57 SVYRCLTFPENISGVETIHLGHTNEEVYKTKGRPENAVMISNQSVDWIIIPMLAARG 116  
QY 111 ALGH---VRYLKEGLKMLPLYGCFYAOHGCIYKRSKAFNEKEMRNKLOSVDAGTPMY 167  
Db 117 DQGNBOAFRVMYKNSIHLVPMFGWYLFQHGCIYVRRFGEFIGAPVROLKMLINESDPY 176

QY 168 LVIFPEGTRVPEQTKVLSAQAFPAORGLAVLKHVLTLPRIKATHVAFDCMKNYLDAIYD 227  
Db 177 LILFPEGTRSAKKHLLLESNNRFLKSGROPONVILCPRSSGLQALDLNLT-IDAITYD 235  
QY 228 VTVVY-EGKRDG-----QRRSEPTWTFIC--KECPRIHIDRIDK 267  
Db 236 VTVWYGMRFDLNLNLTHYLLIYMAERRLAGMFMFCGSOOFKOLHILDRIP 295  
QY 268 KDVEPEOEHRMRWLHERFEIKDKMLIEFESDPERRKRPFGKSVNSKTSIKKTLPSML 327  
Db 296 DEVPKAKLELRTWTIEFTKERRIDIEFY-SEKPSGSGALP-----CPISTLSTLFR 348  
QY 328 LS-GLTAGMLMTDAGRLVNTWYTGTLGCLWTTIK 363  
Db 349 FSAALLAPFRSRTIGR-IYLLTIASSPL-IAMLHIR 383

## RESULT 5

Q9SSH0 PRELIMINARY: PRT: 373 AA.  
AC Q9SSH0: 01-MAY-2000 (TrEMBLrel. 13, Created)  
DT 01-MAY-2000 (TrEMBLrel. 13, last sequence update)  
DT 01-JUN-2001 (TrEMBLrel. 17, last annotation update)  
DE F25A4.2 PROTEIN.  
GN F25A4.2.  
OS Arabidopsis thaliana (Mouse-ear cress).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.  
OX NCBI\_TaxID=3702;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=CV. COLUMBIA;  
RA Vysotskaia V.S., Schwartz J.R., Yu G., Toriumi M., Lenz C., Liu S.,  
Lee J.M., Li J., Gonzalez A., Liu A., Liu K., Vaysberg M., Sakano H.,  
Chin C., Choi E., Chou E., Altafi H., Araujo R., Brooks S.,  
Buehler E., Chao Q., Conn L., Conway A.B., Dunn P., Hansen N.,  
Hwang B., Huizar L., Khan S., Kim C., Palm C., Kowley D., Shinn P.,  
Walker M., Davis R.W., Ecker J.R., Federici N.A., Theologis A.;  
RL "Arabidopsis thaliana chromosome 1 BAC F25A4 sequence.";  
DR EMBL: AC008263; AAD55275.1; -  
DR InterPro: IPR002123; Acyltransferase.  
DR Pfam: PF01553; Acyltransferase; 1.  
SQ SEQUENCE 373 AA: 42383 MW: 98642948D94709E6 CRC64;

Query Match 20.7%; Score 398.5; DB 10; Length 373;  
Best Local Similarity 29.8%; Pred. No. 4.7e-28;  
Matches 97; Conservative 68; Mismatches 117; Indels 43; Gaps 7;

QY 12 MRYLPSVVLGTAPTYVLAMGVWRLLSAFLPARFYQALDDRLTYCYQSMYLT----- 63  
Db 23 LRGLMILVLPVLSAFMFLT-----YFAP---IALGLRLISVQOSKRVSLIFGLM 70  
QY 64 -----FFPENYTGVOILLIGDLPKKNENIYIYLANHSTVDWIVADILAIKRONLGHVRV 118  
Db 71 LALMPLLFETVNTGTVVFSGDILPVEKRVLLIANHRTVEDWMLMIALRKGLGYIKYV 130  
QY 119 LRKGLKMLPLYGCFYAOHGCIYKRSKAFNEKEMRNKLOSVDAGTPMYLIFPEGTRV 178  
Db 131 LKSSLMKLPFGCFVLETPYERKREVDPEVLMLSSFKPQEPMLALFPEGTDIT 190  
QY 179 PEQTKVLSAQAFPAORGLAVLKHVLTLPRIKATHVAFDCMKNYLDAIYDVTVVYEGKRDG 238  
Db 191 EKKCK--RSQKFAAELGVALSNVLLPRTGRGVCLVYLNHSLDAVYDITAYKPR--- 244  
QY 239 GORRESPTWTFICKECP-KIHIDRIDKDYDEOEHRMRWLHERFEIKDKMLIEFYE 297  
Db 245 -----CPSEMDNVFGTDPSEVHLHVRVLLKEIPANEASSAWLMDSFKDKLISDF-- 297  
QY 298 SPDEPKRRKRPFGKSVNSKTSIKKTL 322

Db 298 ----NAQKFPNORPEELSVLKCI 318

RESULT 6  
ID Q9C9P8 PRELIMINARY; PRT; 393 AA.  
AC Q9C9P8; 01-JUN-2001 (TREMBlrel.. 17, Created)  
DT 01-JUN-2001 (TREMBlrel.. 17, Last sequence update)  
DE PUTATIVE ACYL-COA:1-ACYLGlycerol-3-PHOSPHATE ACYLTRANSFERASE, 31588-29381  
GN F9E10.13  
OS Arabidopsis thaliana (Mouse-ear cress).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.  
OX NCBI\_TaxID=3702;

RA [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-CV. COLUMBIA;  
RX MEDLINE=21016719; PubMed=11130712;  
RA Theologis A., Ecker J.F., Palm C.J., Federspiel N.A., Kaul S.,  
Buehler E., Chan A., Chao Q., Chen H., Cheuk R.F., Chin C.W.,  
Chung M.K., Conn L., Conway A.B., Conway A.R., Creasy T.H., Dewar K.,  
Dunn P., Elgu P., Feldblum T.V., Feng J.-D., Fong B., Fujii C.Y.,  
Gall J.E., Goldsmith A.C., Haas B., Hansen N.F., Hughes B., Hultzer L.,  
Hunter J.L., Jenkins J., Johnson-Hopson C., Khan S., Khaykin E.,  
Kim C.J., Koo H.L., Kremetska I., Kurtz D.B., Kwan A., Lam B.,  
Langin-Hooper S., Lee A., Lee J.M., Lenz C.A., Li J.H., Li Y.-P.,  
Lin X., Liu S.X., Liu Z.A., Lueros J.S., Maiti R., Marziani A.,  
Millscher J., Miranda M., Nguyen M., Nierman W.C., Osborne B.I.,  
Pai G., Peterson J., Pham P.K., Rizzo M., Rooney T., Rowley D.,  
Sakano H., Salerg S.L., Schwartz J.R., Shinn P., Southwick A.M.,  
Sun H., Tallon L.J., Tambunga G., Tortum M.J., Town C.D., Walker M.,  
Utterback T., Van Aken S., Vaysberg M., Vysotskaya V.S., Walker M.,  
Wu D., Yu G., Fraser C.M., Venter J.C., Davis R.W.;  
RT "Sequence and analysis of chromosome 1 of the plant Arabidopsis  
thaliana."  
RL Nature 408:816-820(2000).  
EMBL: AC013258; MAG51931.1; -  
DR InterPro: IPR002123; Acyltransferase.  
DR Pfam: PF01553; Acyltransferase; 1.  
KW Transferase; Acyltransferase.  
SQ SEQUENCE 393 AA; 44892 MW; FA5949D21E8E9A29 CRC64;

Query Match 20.7%; Score 398.5; DB 10; Length 393;  
Best Local Similarity 29.8%; Pred. No. 5.1e-28;  
Matches 97; Conservative 68; Mismatches 117; Indels 43; Gaps 7;

Db 12 MRYLPSVLLGTPRYVLANGVRLSAFLPAFYQALDRLCYQSVNL----- 63  
Db 23 LRGIMLVLELSTAFMFL-----YFAP---IALGLRLSVQSRKVVSLIGLW 70  
Qy 64 -----PFENVYGOVILLYGCLPKENKIITYLANHOSVDMIVADIAIRONALGHRYV 118  
Db 71 LALMPYLFEEVNGTVVFSDDIIPVEKRVLLIANHRTEDVMYIMNALRKGCGIYKIY 130  
Qy 119 LKESGLKPLRYGCFYFAOHGCIYKRSKAFNEKEMRNKLOSYVDAGTPMYLVIPREGTRYN 178  
Db 131 LKSLMLPLTFGWFHLEFIYPERKREVDENVLLQMLSSFKDQDEPLMLALFPEGTDFI 190  
Qy 179 PEQTKVLSAQAFAAAGLAVLKHVLPRIKATYVAFDCAKNYIDALYDVTVYEGDDG 238  
Db 191 EEKCK---RSQKFAAEVGLPALSIVLLPKTRGEGVCLVILANSIDAVYDLTIAKPR-- 244  
Qy 239 GORREPTMEFLCKECP-KIHIIHDIRDKDVEDEQEHRRMLHEFEIKDKMLIFEYE 297  
Db 245 -----CSPFDNVYFGIDPSEVHIHVRVLLKELIPANEASSAMLMDSFKLKDLSDF-- 297

Qy 298 SPDEPRKRRPFGKSVNSKLSIKKTL 322  
Db 298 ----NAQKFPNORPEELSVLKCI 318

RESULT 7  
ID Q9LHN4 PRELIMINARY; PRT; 375 AA.  
AC Q9LHN4; 01-OCT-2000 (TREMBlrel.. 15, Created)  
DT 01-OCT-2000 (TREMBlrel.. 15, Last sequence update)  
DE 01-DEC-2001 (TREMBlrel.. 19, Last annotation update)  
DE ACYL-COA:1-ACYLGlycerol-3-PHOSPHATE ACYLTRANSFERASE.  
OS Arabidopsis thaliana (Mouse-ear cress).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.  
OX NCBI\_TaxID=3702;

RA [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-COLUMBIA;  
RA Kaneko T., Kato T., Sato S., Nakamura Y., Asamizu E., Tabata S.;  
RL Submitted (May-2000) to the EMBL/GenBank/DBJ databases.  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN-COLUMBIA;  
RX MEDLINE=20363099; PubMed=10907853;  
RA Nakamura Y.;  
RT "Structural analysis of Arabidopsis thaliana chromosome 3. II.  
RT TNC and BAC clones."  
RL DNA Res. 7:217-221(2000).  
EMBL: AP002039; BAB03094.1; -  
DR InterPro: IPR002123; Acyltransferase.  
DR Pfam: PF01553; Acyltransferase; 1.  
KW Transferase; Acyltransferase.  
SQ SEQUENCE 375 AA; 43490 MW; 8F6F5DC313A3D704 CRC64;

Query Match 20.3%; Score 390.5; DB 10; Length 375;  
Best Local Similarity 26.9%; Pred. No. 2.6e-27;  
Matches 101; Conservative 78; Mismatches 128; Indels 69; Gaps 9;

Qy 12 MRYLPSVLLGTPRYVLANGVRLSAFLPAFYQALDRLCYQSVNLFFF 66  
Db 18 LRGICIMLVLELSTAFMFLIFWGLSAVLRIFS---IRYSRKCVSFFGSMALMPFLF 73  
Qy 67 ENTGVOILLGDLPRKKNKIITYLANHOSVDMIVADIAIRONALGHRYVLEKGLKWL 126  
Db 74 EKINKTVITSGDKVPCEDVLLIANHRTEDVMYIMNALRKGQIGNITVLSKSLML 133  
Qy 127 PLYGCFYFAOHGCIYKRSKAFNEKEMRNKLOSYVDAGTPMYLVIPREGTRYNPEQTKVLS 186  
Db 134 PLGMAHLEEFIPVERKREVDENVLLQYISFRKDRPDALMLALFPEGTITYTAQCO--- 190  
Qy 187 ASQAFAAAGLAVLKHVLPRIKATYVAFDCAKNYIDALYDVTVYEGKDDGOGORRESPT 246  
Db 191 RSKKFAENCLPIILNVLPRTFGFSCLQELSCSDAVYDVITGYKTR----- 239  
Qy 247 MTEFLCKECP-----KIHIIHDIRDKDVEDEQEHRRMLHEFEIKDKMLIFE 234  
Db 240 -----CSPFDNVYGIPESEVHIHRIINILQIPNOBKIDNAMLNMTFOLKDLND 291  
Qy 295 FYESPDERKRRPFGKSVNSKLSIKKTLPLMLISGLT-----AGMLMTDAGRKLY 345  
Db 292 FYSN-----GHRPNEGTEKEFPYTKKTLNCLAVIAFTTICTHLPFPSSIMWF-----RIT 341  
Qy 346 VNTWYIGTLGCLMVT 361  
Db 342 VS-----LACYVLT 350

RESULT 8

Q41745 Q41745 PRELIMINARY; PRT; 374 AA.  
 AC Q41745:  
 DT 01-NOV-1996 (TREMblrel. 01, Created)  
 DT 01-NOV-1996 (TREMblrel. 01, last sequence update)  
 DT 01-DEC-2001 (TREMblrel. 19, last annotation update)  
 DE 1-ACYL-GLYCEROL-3-PHOSPHATE ACYLTRANSFERASE (PUTATIVE).  
 OS Zea mays (Maize).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC clade;  
 OC Panicoideae; Andropogoneae; Zea.  
 OX NCBI\_TaxID=4577;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-BLACK MEXICAN SWEET; TISSUE-ENDOSPERM;  
 RX MEDLINE=95035993; PubMed=7948871;  
 RA Brown A.P., Coleman J., Tomney A.M., Watson M.D., Slabas A.R.;  
 RT "Isolation and characterization of a maize cDNA that complements a 1-  
 RT acyl-sn-glycerol-3-phosphate acyltransferase mutant of E. coli and  
 RT encodes a protein which has similarities to other acyltransferases.";  
 RL Plant Mol. Biol. 26:211-223(1994).  
 DR EMBL; Z29518; CAA82638.1; -  
 DR InterPro; IPR002123; Acyltransferase.  
 DR InterPro; IPR003015; HLH\_Myc.  
 DR Pfam; PF01553; Acyltransferase; 1.  
 DR PROSITE; PS00038; HELIX\_LOOP\_HELIX; UNKNOWN\_1.  
 KW Transferase; Acyltransferase.  
 SQ SEQUENCE 374 AA; 42571 MW; F1F5492CAFF24F93 CRC64;

Query Match 17.0%; Score 328; DB 10; Length 374;  
 Best Local Similarity 29.8%; Pred. No. 1,3e-21;  
 Matches 92; Conservative 64; Mismatches 121; Indels 32; Gaps 8;

QY 29 VLAMGVMLLSA-----FLPAR-----FYQALDRLYCYQSMVLEFFENTYQV 73  
 DB 9 VLPLLEFLISGLINVAIOAVLEVTIRPFSKSFYRINRFLMLQLVWVVMMAVYK 68  
 QY 74 ILLYD-----LPKNENIITYLANHOSTVDVIADILAIKONLGHVRYLKEGLKWLPLX 129  
 DB 69 VOLHDEETYSRSMGEHALIISNHRSDIDWLIGWILAQSGCLGSTLAVMKSSKFLPYI 128  
 QY 130 G--CYFAOHGIIYKRSKAFNEKEMRNKLOSVDAGTPMYLVIPEEGTRVNEQTRVLSA 187  
 DB 129 GMSMFAEY--LELERSNAKDEKTLKMGQLKDEPRFWLALFVEGRTFTPAK---LLA 183  
 QY 188 SOAFPAQKGLAVLKHVLPRIKATHVAFDCMKNYLDATVTVYEGKDDGQRRSPPTM 247  
 DB 184 AQEYAAOSGLPAPRVNLLPRTKGFVSASVIMRDFVPAIDTVIYP-KDS-----PQPTM 237  
 QY 248 TFFLCKECPKTHIHIDRIKDVPEQEHMRRLHERFEIKMKMLIEFESPDPERKRF 307  
 DB 238 LRLGQSSVTHVRMKRHAMSEMPKSDSDVSKWCKDIFAKALLDKHLATGTFTDEIRP 297  
 QY 308 GKSVNSKL 316  
 DB 298 IGRPVKSL 306  
 RESULT 9  
 Q9XFW4 Q9XFW4 PRELIMINARY; PRT; 390 AA.  
 AC Q9XFW4:  
 DT 01-NOV-1999 (TREMblrel. 12, Created)  
 DT 01-NOV-1999 (TREMblrel. 12, last sequence update)  
 DT 01-DEC-2001 (TREMblrel. 19, last annotation update)  
 DE ACYL-COA:1-ACYLGLYCEROL-3-PHOSPHATE ACYLTRANSFERASE (EC  
 DE 2.3.1.51).  
 GN LPAAT.  
 OS Brassica napus (Rape).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
 OC eurosids II; Brassicales; Brassicaceae; Brassica.

OX NCBI\_TaxID=3708;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE-DEVELOPING SILIQUES;  
 RA Graefin zu Muenster A., Wolter F.P., Frentzen M.;  
 RT "A cDNA encoding a microsomal 1-acylglycerol-3-phosphate  
 RT acyltransferase of Brassica napus L.";  
 RL Submitted (MAY-1997) to the EMBL/Genbank/DBJ databases.  
 DR EMBL; Z295637; CAB09138.1; -  
 DR InterPro; IPR002123; Acyltransferase.  
 DR InterPro; IPR003015; HLH\_Myc.  
 DR Pfam; PF01553; Acyltransferase; 1.  
 DR PROSITE; PS00038; HELIX\_LOOP\_HELIX; UNKNOWN\_1.  
 KW Transferase; Acyltransferase.  
 SQ SEQUENCE 390 AA; 43771 MW; F1446ELB3009C37 CRC64;

Query Match 16.1%; Score 309.5; DB 10; Length 390;  
 Best Local Similarity 26.5%; Pred. No. 6,8e-20;  
 Matches 83; Conservative 75; Mismatches 118; Indels 37; Gaps 9;

QY 69 YRGVQILLYGD---LPKNENIITYLANHOSTVDVIADILAIKONLGHVRYLKEGLK 124  
 DB 63 WAGVKIQVFADETFENRMGKEHALVYCNHRSDIDWLIGWILAQSGCLGSALAVMKSSK 122  
 QY 125 WPLVIG--CYFAOHGIIYKRSKAFNEKEMRNKLOSVDAGTPMYLVIPEEGTRVNEQ 182  
 DB 123 FLPLVIGSMVFSEY--LELERNMAKDESTLQSGLORLNDFPRFWLALFVEGRTFTPAK- 179  
 QY 183 KVLASQAFPAQKGLAVLKHVLPRIKATHVAFDCMKNYLDATVTVYEGKDDGQRR 242  
 DB 180 --LKAQGEYAAASELPVRPNVLLPRTKGFVSASVIMRDFVPAIDTVIYP-----KTS 231  
 QY 243 EEPYTFELCKCPKTHIHIDRIKDVPEQEHMRRLHERFEIKMKMLIEFESPDPE 302  
 DB 232 PPPTMLRLFGQPSVYVHVHAIKCHSMKDLPEPEDEIAQWCRDQFVAKDALDKHIAA---- 287  
 QY 303 RRRPRGKSVNSKLSIKTKLPSMLIISG---LTAGML-----MTDGRKLYVNTWYI 351  
 DB 288 --DTFPGQ--KEQNGRPKSLAVVSWACLLTJGMAFLHWSNLFSSKKAIASAFGL 342  
 QY 352 GTLLGCLWYTIKA 364  
 DB 343 GIITLCMOILINS 355

RESULT 10  
 Q40119 Q40119 PRELIMINARY; PRT; 377 AA.  
 AC Q40119:  
 DT 01-NOV-1996 (TREMblrel. 01, Created)  
 DT 01-NOV-1996 (TREMblrel. 01, last sequence update)  
 DT 01-DEC-2001 (TREMblrel. 19, last annotation update)  
 DE 1-ACYL-SN-GLYCEROL-3-PHOSPHATE ACYLTRANSFERASE (PUTATIVE).  
 OS Limnathes douglasii (Douglas's meadowfoam).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
 OC eurosids II; Brassicales; Limnathaceae; Limnathes.  
 OX NCBI\_TaxID=28973;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC MEDLINE=96046746; PubMed=7579178;  
 RA Brown A.P., Brough C.L., Kroon J., Slabas A.R.;  
 RT "Identification of a cDNA that encodes a 1-acyl-sn-glycerol-3-  
 RT phosphate acyltransferase from Limnathes douglasii.";  
 RL Plant Mol. Biol. 29:267-278(1995).  
 DR EMBL; Z48730; CAA88620.1; -  
 DR InterPro; IPR002123; Acyltransferase.  
 DR InterPro; IPR003015; HLH\_Myc.  
 DR Pfam; PF01553; Acyltransferase; 1.  
 DR PROSITE; PS00038; HELIX\_LOOP\_HELIX; UNKNOWN\_1.  
 KW Transferase; Acyltransferase.  
 SQ SEQUENCE 377 AA; 42780 MW; 0DECDEF25D39687C CRC64;

```
Query Match      16.0%; Score 307; DB 10; Length 377;
Best Local Similarity 25.5%; Pred. No. 1,1e-19;
Matches 91; Conservative 85; Mismatches 131; Indels 50; Gaps 13;

QY 37 LLSAFLPARRY---QALDDRLCYCYQSMV-----LFFEFENTGQIILIGDLPR--- 82
DB 21 LVNFIQAVFEYVLLVPRISKDTYRRINTLVALLMLLVWIDWAGKAVOLYDTDESFR 80
QY 83 -NKENIYILNHOSTVMIADILAIQNALGHWYLVKLGKMLPLYG--CYFAOHGHI 139
DB 81 MGKEHALLICNHRSDIDMLGVWLAQRCGLSSIAVMKSSKFLPYIGMSMWSEY--L 138
QY 140 YKRSAKFNEKEMRNKLSQSYVDAQTPMYLVIFPEGTRYNEPQTKV-LSASQAFPAQGLA 198
DB 139 FLENNMAKDENTLKSGLQRLNDPFRFLALFVEGTRF---TRAKULDAQEVNAGSLP 194
QY 199 VLKHVLTFRKATIVAA-DCCKKNVLDIAYTVVYEGKDDGGQRRSESTMTMEFLCKECPKI 258
DB 195 VPRNVLLPRTKGFVSANVSNMRSFPAIYDLTVLRP-----KTTEQPTMLRLEFRKSSVY 248
QY 259 HIHDIRDKDQVPEQIHMNRMLHREFEIKDKMLIEFYESPDPERRKRPFGKSVNSKLSI 318
DB 249 HVHLKRLMKDLPRKTDICVAKQCKDPTISKD-ALLDKHVAED-----TFSGLEYQ---DI 299
QY 319 KTLPLSMILISG-----LTAGML-----MTDAGRKLVTWTWYIGLLGLCLWYTIKA 364
DB 300 GRPMKSLVWVWVMWMLCLGLGVKRLQMSALLSSWKGMITTFVGLGYTVLMLILRS 356

RESULT 11
Q9SDN3 PRELIMINARY; PRT; 306 AA.
AC Q9SDN3;
DT 01-MAY-2000 (TReMBLrel. 13, Last sequence update)
DT 01-DEC-2001 (TReMBLrel. 19, Last annotation update)
DE 1-ACYL-SN-GLYCEROL-3-PHOSPHATE ACYLTRANSFERASE.
OS Prunus dulcis (Almond) (Prunus amygdalus).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids I; Rosales; Rosaceae; Amygdaloideae; Prunus.
OX NCBI_TaxID=3735;
RN SEQUENCE FROM N.A.
RA Campalans A., Pages M., Messegueur R.;
RT Identification of differentially expressed genes during dehydration
RT in Almond (Prunus amygdalus) using the cDNA-AFLP technique."
RL EMBL: AF213937; AAF20003.1;
DR EMBL: AF213937; AAF20003.1;
DR InterPro: IPR002123; Acyltransferase.
DR InterPro: IPR003015; HLH_Myc.
DR Pfam: PF01553; Acyltransferase; 1.
DR PROSITE: PS00038; HELIX_LOOP_HELIX; UNKNOWN_1.
DR Transferrase; Acyltransferase.
KW TRANSFERRASE; ACYLTRANSFERASE.
SQ SEQUENCE 306 AA; 33993 MW; CBD334496E25908D CRC64;
```

```
Query Match      15.9%; Score 305.5; DB 10; Length 306;
Best Local Similarity 31.9%; Pred. No. 1,1e-19;
Matches 80; Conservative 55; Mismatches 101; Indels 15; Gaps 7;

QY 84 KENIYILNHOSTVMIADILAIQNALGHWYLVKLGKMLPLYG--CYFAOHGSIYV 141
DB 3 KENALVLSNHRSDIDMLGVWLAQRCGLSSIAVMKSSKFLPYIGMSMWSEY--LFL 60
QY 142 KRSAKFNEKEMRNKLSQSYVDAQTPMYLVIFPEGTRYNEPQTKV-LSASQAFPAQGLA 201
DB 61 ESRWAKDEGTLKSGVGRFLKDPQFWMFLALFVEGTRF--QAKLL-AAQEVNAGSLPVR 117
QY 202 HVLTFRKATIVAA-DCCKKNVLDIAYTVVYEGKDDGGQRRSESTMTMEFLCKECPKI 261
DB 231 AMACLLTIGAL 241

RESULT 12
Q9VW51 PRELIMINARY; PRT; 386 AA.
AC Q9VW51; Q9VW50;
DT 01-MAY-2000 (TReMBLrel. 13, Created)
DT 01-MAY-2000 (TReMBLrel. 13, Last sequence update)
DT 01-JUN-2001 (TReMBLrel. 17, Last annotation update)
DE CC4729 PROTEIN.
GN CC4729.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Phryganea; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Phryganeidae; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN SEQUENCE FROM N.A. (LONG AND SHORT ISOFORMS).
RP STRAIN=BERKELEY;
RC MEDLINE=20196006; PubMed=10731132;
RX ADAMS M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazer R.G., Champe M., Pfeiffer B.D.,
RA Wen K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abril J.F., Aghayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Bertone P., Bhandari D., Bolintsov S.,
RA Borokova D., Botchan M.R., Bouck J., Brooksstein P., Brotler P.,
RA Burris K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson R., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Fostler C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Idegami C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Moberly C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusslein D.R., Paclet J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Sytkas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissenbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster."
RL Science 287:2185-2195(2000).
CC -1- ALTERNATIVE PRODUCTS: 2 ISOFORMS; A LONG FORM (SHOWN HERE) AND A
CC SHORT FORM; ARE PRODUCED BY ALTERNATIVE SPLICING.
DR EMBL: AE003527; AAF49471.1;
DR EMBL: AE003527; AAF49472.1;
DR FlyBase; FBgn0036623; CG4729.
DR InterPro: IPR002123; Acyltransferase.
```

Query Match	13.9%;	Score 267;	DB 5;	Length 380;
Best Local Similarly	25.9%;	Pred. No. 5e-16;		
Matches 83;	Conservative	69;	Mismatches 123;	Indels 46;
				Gaps 9



GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 28, 2002, 11:18:44 ; Search time 75.33 Seconds

(without alignments)  
336.185 Million cell updates/sec

Title: US-09-853-526-70

Perfect score: 228

Sequence: 1 MRYLLPSVVLGTAFTYVLA.....NYLDAIYDVWVVEKDDGG 228

Scoring table: OLIGO  
Gapop 60.0 , Gapext 60.0

Searched: 747574 seqs, 11073796 residues

Word size : 0

Total number of hits satisfying chosen parameters: 747574

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database :

A.Geneseq\_032802:\*  
1: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA1980.DAT:\*  
2: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA1981.DAT:\*  
3: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA1982.DAT:\*  
4: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA1983.DAT:\*  
5: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA1984.DAT:\*  
6: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA1985.DAT:\*  
7: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA1986.DAT:\*  
8: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA1987.DAT:\*  
9: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA1988.DAT:\*  
10: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA1989.DAT:\*  
11: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA1990.DAT:\*  
12: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA1991.DAT:\*  
13: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA1992.DAT:\*  
14: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA1993.DAT:\*  
15: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA1994.DAT:\*  
16: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA1995.DAT:\*  
17: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA1996.DAT:\*  
18: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA1997.DAT:\*  
19: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA1998.DAT:\*  
20: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA1999.DAT:\*  
21: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA2000.DAT:\*  
22: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA2001.DAT:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	228	100.0	228	20	AAV36740	Truncated PGI prot
2	228	100.0	352	20	AAV36728	Human PGI protein
3	228	100.0	353	22	AAAB3591	Human polypeptide
4	228	100.0	353	22	AAAB85532	Human secreted pro
5	228	100.0	353	22	AAAB85552	Human secreted pro
6	228	100.0	364	20	AAV36729	Human PGI protein
7	228	100.0	364	22	AAAM50128	Human acyltransfer
8	228	100.0	364	22	AAU00665	Human lysophosphat
9	185	81.1	185	20	AAV36753	PGI splice variant
10	185	81.1	315	20	AAV36751	PGI splice variant
11	184	80.7	300	20	AAV36752	PGI splice variant

12	154	67.5	182	20	AAV36750	PGI splice variant
13	144	63.2	269	22	AAE09653	Human gene 15 enco
14	144	63.2	353	22	AAE09655	Human protein sequ
15	127	55.7	372	22	AAAM41377	Human polypeptide
16	105	46.1	291	20	AAV36744	PGI splice variant
17	85	37.3	96	20	AAV36749	PGI splice variant
18	74	32.5	261	20	AAV36745	PGI splice variant
19	70	30.7	257	22	ABG01360	Novel human diagno
20	70	30.7	980	22	ABG11977	Novel human diagno
21	62	27.2	66	20	AAV36748	PGI splice variant
22	62	27.2	68	20	AAV36747	PGI splice variant
23	62	27.2	77	20	AAV36742	PGI splice variant
24	62	27.2	90	20	AAV36746	PGI splice variant
25	62	27.2	238	20	AAV36743	PGI splice variant
26	44	19.3	97	22	ABG12868	Novel human diagno
27	44	19.3	1104	22	ABG12107	Novel human diagno
28	32	14.0	354	20	AAV36741	Mouse PGI protein
29	9	3.9	9	20	AAV36734	PGI box 3 motif.
30	9	3.9	157	22	AAAM50130	Lysophosphatidic a
31	9	3.9	237	18	AAV55540	H. pylori ORF 12ge
32	9	3.9	240	19	AAV98477	H. pylori GHPD 903
33	9	3.9	242	18	AAV55441	H. pylori ORF 06gp
34	8	3.5	9	20	AAV36737	Acyl glycerol tran
35	8	3.5	239	22	ABBS2710	Escherichia coli p
36	8	3.5	396	21	AAV99484	Yeast acyltransfer
37	8	3.5	1820	22	ABBS2370	Drosophila melanog
38	7	3.1	12	16	AAV87705	Cocount AG3PAT try
39	7	3.1	12	20	AAV43149	Cocount 1-acylglyc
40	7	3.1	65	16	AAV87720	Cocount LPAAT clon
41	7	3.1	65	20	AAV43164	1-acylglycerol -3-p
42	7	3.1	82	16	AAV87719	Cocount LPAAT clon
43	7	3.1	82	16	AAV87719	Cocount LPAAT clon
44	7	3.1	82	20	AAV43160	1-acylglycerol -3-p
45	7	3.1	82	20	AAV43163	1-acylglycerol -3-p

#### ALIGNMENTS

RESULT 1	
AAV36740	standard; Protein; 228 AA.
XX	
AC	AAV36740;
XX	
DT	27-SEP-1999 (first entry)
XX	
DE	Truncated PGI protein sequence.
XX	
KW	PGI gene; biallelic marker; human; PSA; PGI-related biallelic marker;
KW	cancer; prostate cancer; diagnosis; therapy; prostate specific antigen.
XX	
OS	Homo sapiens.
XX	
PN	W0932644-A2.
XX	
PD	01-JUL-1999.
XX	
PF	22-DEC-1998; 98WO-1B02133.
XX	
PR	09-SEP-1998; 98US-009658.
XX	
PR	22-DEC-1997; 97US-0996306.
XX	
PA	(GEST) GENSET.
XX	
PI	Blumenfeld M, Bougueleret L, Chumakov I, Cohen D;
XX	WPI; 1999-405178/34.
XX	N-PSDB; AA00926.
DR	
XX	
PT	Use of a prostate cancer associated gene and biallelic markers
XX	derived from it

PS Claim 7; Page 208; 385pp; English.

The invention relates to a mammalian Pcl gene and protein, and a set of Pcl biallelic markers. The pcl polynucleotide and biallelic markers are used in a hybridisation assay, a sequencing assay, or in an allele-specific amplification assay for determining the identity of a nucleotide at a pcl-related biallelic marker. The methods can be used to detect and to assess the risk of developing cancer or prostate cancer. Early-stage diagnosis of prostate cancer relies on prostate specific antigen (PSA) dosage. However, the effectiveness of this is limited due to its inability to discriminate between malignant and non-malignant affections of the organ. A need exists for both a reliable diagnostic procedure which would enable early-stage diagnosis, and for preventative and curative treatments of the disease. The Pcl gene can be used for detection of prostate cancer, and the risk of developing it in the future, and can also be used to determine therapies for the disease.

**SQ Sequence 228 AA;**

Query Match	100.0%;	Score 228;	DB 20;	Length 228;
Best Local Similarity	100.0%;	Pred. No. 2.1e-222;		
Matches 228;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;

Qy	1	MRLLPSVLLGTAPTAYLAWGVR	SSAF	PAFFQALDRLCYVOSMFLFEFENYTG	60
Db	1	mcyllpsvlllgtpacpçylawgwrl	lsaf	parifgaldrlçycvgsmlffifemtyç	60
Qy	61	VOILLGDLPRKKNENIIVLANHGQSTVDIM	IVADII	AIRONALGHNRYVLKEGLKMLPLVGC	120
Db	61	vqlllgdglprknellilylanhgstvdwlvad	lla	iqnalgntrryylkeglkklpbygc	120
Qy	121	YFAOHGGLYKRSAAFNKEKEMRNKLQSYVDAGT	PMYLV	IFEEGSTRNDEQTRVLASQAFL	180
Db	121	yfaohgglçykrçakfnekemrnlqsyvdag	tçmylv	lfefçitrypçeqçklvstsqaf	180
Qy	181	AAQKGLANLVKHIIVPRIRATIVAADCKMNI	IDAL	IVDTIVYIEGKDDG	228
Db	181	aagçglavlkhlvlpçrlaktvadçkmnyldal	dydv	çvyvçgçddçg	228

## RESULT 2

ID AAY36728 standard; Protein; 352 AA.

AC AAY36728;

DT 27-SEP-1999 (first entry)

DE Human PGI<sub>2</sub> protein sequence.

KM PGI gene; biallelic marker; human; PSA; PGI-related biallelic marker;  
KM cancer; prostate cancer; diagnosis; therapy; prostate specific antigen;

05 *Homo sapiens*.

PN W09932644-A2.

PD 01-JUL-1999.

22-DEC-1998; 98WO-IB02133.

PR 09-SEP-1998; 98US-0099658.

XX

XX

[illegible]

DR N-PSDB; AA200872.

PT	Use of a prostate
1	1
2	1
3	1
4	1
5	1
6	1
7	1
8	1
9	1
10	1
11	1
12	1
13	1
14	1
15	1
16	1
17	1
18	1
19	1
20	1
21	1
22	1
23	1
24	1
25	1
26	1
27	1
28	1
29	1
30	1
31	1
32	1
33	1
34	1
35	1
36	1
37	1
38	1
39	1
40	1
41	1
42	1
43	1
44	1
45	1
46	1
47	1
48	1
49	1
50	1
51	1
52	1
53	1
54	1
55	1
56	1
57	1
58	1
59	1
60	1
61	1
62	1
63	1
64	1
65	1
66	1
67	1
68	1
69	1
70	1
71	1
72	1
73	1
74	1
75	1
76	1
77	1
78	1
79	1
80	1
81	1
82	1
83	1
84	1
85	1
86	1
87	1
88	1
89	1
90	1
91	1
92	1
93	1
94	1
95	1
96	1
97	1
98	1
99	1
100	1

PT	derived from it
XX	
PS	Claim 7; Page 190; 385pp; English

The invention relates to a mammalian PGI gene and protein, and a set of PGI biallelic markers. The PGI polynucleotide and biallelic markers are used in a hybridisation assay, a sequencing assay, or in an allele-specific amplification assay for determining the identity of a nucleotide at a PGI-related biallelic marker. The methods can be used to detect and to assess the risk of developing cancer or prostate cancer. Early-stage diagnosis of prostate cancer relies on prostate specific antigen (PSA) dosage. However, the effectiveness of this is limited due to its inability to discriminate between malignant and non-malignant affections of the organ. A need exists for both a reliable diagnostic procedure which would enable early-stage diagnosis, and for preventative and curative treatments of the disease. The PGI gene can be used for detection of prostate cancer, and the risk of developing it in the future, and can also be used to determine therapies for the disease.

Sequence 352 AA;

Query Match	100.0%;	Score 228;	DB 20;	Length 352;
Best Local Similarity	100.0%;	Pred. No. 3.1e-222;		
Matches 228;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;

Qy	1	MRYLPSVVLGGTAPTYVLAAGVWRLSAFLPAFHYQALDRLCVYQSMVLEFFENYTG	60
Db	1	mrYllpsvvlGGtApTyVlaAGvWRlSAfLPaFHYqALdRlCVyQsMvLEffENyTg	60
Qy	61	VOILLGDLPKNKENIITYLANHOSVPMIVADILAIROMALGHRYVYLKEGILMKLPLRYGC	120
Db	61	vGilllgdLpKnKenIiTyLaNgStVdVadIlaiRnaLghRyVylKegIlkLpLryGc	120
Qy	121	YFAOHGGIYVRSAKFENEKEMRNKLQSYVDAGTMYLVIPEEGTRYNEQOTKVLASQAF	180
Db	121	yfaOhGgiYvRsaKfEnEkEmrNkLqSyVdAGtMyLviPeEGTrYNeQoTKvLAsQaf	180
Qy	181	AAQGRVLAVLKHVLPRIKATIVAADCKMKNYIDATYDTVYVEGDDSG	228
Db	181	aAQgRvLAvLKhvLPriKaTiVaADcKMNyIdATyDtVYvEGdDsg	228

### RESULT 3

ID	AAM39591	standard; Protein; 353 AA.
----	----------	----------------------------

AC AAM39591;

DT 22-OCT-2001 (first entry)

Human polypeptide SEQ ID NO 2736.

KM Human; nootropic; immunosuppressant; cytostatic; gene therapy; cancer;  
KW peripheral nervous system; neuropathy; central nervous system; CNS;

KW amyotrophic

KW      leukaemia.

OS Homo sapiens

PN WO200153312-

PD 26-JUL-2001.

PF 26-DEC-2000;

PR 21-JAN-2000;

PR 09-JUL-2000;



PR 14-SEP-2000; 2000US-0662191.  
 PR 19-OCT-2000; 2000US-0693036.  
 PR 29-NOV-2000; 2000US-0727344.  
 XX  
 PA (HYSE-) HYSEQ INC.  
 XX  
 PI Tang YF, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;  
 PI Zhao Q, Zhou P, Goodrich R, Drmanac RT;  
 PI  
 XX WPI: 2001-442253/47.  
 DR N-PSDB; AAI58747.  
 XX  
 PT Novel nucleic acids and polypeptides, useful for treating disorders  
 PT such as central nervous system injuries -  
 XX  
 PS Example 4; SEQ ID NO 2736; 10078bp; English.  
 CC The invention relates to human nucleic acids (AAI57798-AAI61369) and  
 CC the encoded polypeptides (AAM3642-AAM42213) with nootropic,  
 CC immunosuppressant and cytostatic activity. The polynucleotides are useful  
 CC in gene therapy. A composition containing a polypeptide or polynucleotide  
 CC of the invention may be used to treat diseases of the peripheral nervous  
 CC system, such as peripheral nervous injuries, peripheral neuropathy and  
 CC localised neuropathies and central nervous system diseases, such as  
 CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic  
 CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the  
 CC utilisation of the activities such as: immune system suppression,  
 CC Actin/Inhibin activity, chemotactic/chemokinetic activity, haemostatic  
 CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,  
 CC assays for receptor activity, arthritis and inflammation, leukemias and  
 CC C.N.S disorders.  
 CC Note: The sequence data for this patent did not form part of the printed  
 CC specification.  
 CC  
 XX  
 SQ Sequence 353 AA;  
 XX  
 Query Match 100.0%; Score 228; DB 22; Length 353;  
 Best Local Similarity 100.0%; Pred. No. 3.1e-222;  
 Matches 228; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 MRLTSLVVLGTAFTYVLAAGWVRLISAFLPARFOALDDRLCYQSWVLEFFENYTG 60  
 DB 1 MYLLPSVVLGTAFTYVLAAGWVRLISAFLPARFOALDDRLCYQSWVLEFFENYTG 60  
 QY 61 VQILLYGDLPKKNENITILANHOSVDWIVADILAIRONALGHVRYLKEGLKWLPLYGC 120  
 DB 61 VQILLYGDLPKKNENITILANHOSVDWIVADILAIRONALGHVRYLKEGLKWLPLYGC 120  
 QY 121 YFAOHGIIYVKRSKAKFNEKEMRNKLOSVDAGTPMVLVIFPEGTRNPEQTKVLSAQAR 180  
 DB 121 YFAOHGIIYVKRSKAKFNEKEMRNKLOSVDAGTPMVLVIFPEGTRNPEQTKVLSAQAR 180  
 QY 181 AAOAGLAVLKHLVTPRIKATHVAFDCMKNYLDAIVDTVYVEEGRDGG 228  
 DB 181 AAOAGLAVLKHLVTPRIKATHVAFDCMKNYLDAIVDTVYVEEGRDGG 228  
 RESULT 4  
 AAB85532  
 ID AAB85532 standard; protein; 353 AA.  
 XX  
 AC AAB85532;  
 XX  
 DT 25-SEP-2001 (first entry)  
 DE Human secreted protein (clone id HLT084).  
 XX  
 KW Secreted protein; immunosuppressive; antiarthritic; antineumatic;  
 KW antiproliferative; cytostatic; cardiac; vasotropic; cerebroprotective;  
 KW nootropic; neuroprotective; antibacterial; virucide; fungicide; human;  
 KW optical/maleological; gene therapy.

XX  
 OS Homo sapiens.  
 XX  
 PN WO200155430-A1.  
 XX  
 PD 02-AUG-2001.  
 XX  
 PF 17-JAN-2001; 2001WO-US01431.  
 XX  
 PR 31-JAN-2000; 2000US-0179065.  
 PR 04-FEB-2000; 2000US-0180628.  
 PR 12-SEP-2000; 2000US-0231968.  
 XX  
 PA (HUMA-) HUMAN GENOME SCI INC.  
 XX  
 PI Rosen CA, Komatsoulis GA, Baker KP, Birse CE, Soppet DR, Olsen HS;  
 PI Moore PA, Wei P, Ebner R, Duan DR, Shi Y, Choi GH, Fiscella M;  
 PI N J, Ruben SM, Barash SC;  
 XX  
 DR WPI: 2001-476220/51.  
 DR N-PSDB; AAM46942.  
 XX  
 PT 17 Isolated nucleic acid molecules encoding human secreted proteins,  
 PT used to preventing, treating or ameliorating a medical condition -  
 XX  
 PS Claim 11; Page 450-451; 482pp; English.  
 CC The invention provides novel human secreted proteins and polynucleotides  
 CC encoding them. The secreted proteins can be expressed by standard  
 CC recombinant methodology. The secreted proteins and polynucleotides are  
 CC used to prevent, treat or ameliorate a medical condition in e.g. humans,  
 CC mice, rabbits, goats, horses, cats, dogs, chickens or sheep. They can  
 CC also be used in diagnosing a pathological condition. The antibodies to  
 CC the proteins can also be used in alleviating symptoms associated with the  
 CC disorders and in diagnostic immunoassays e.g. radioimmunoassays or enzyme  
 CC linked immunosorbent assays (ELISA). Disorders which are diagnosed or  
 CC treated include autoimmune diseases e.g. rheumatoid arthritis,  
 CC hyperproliferative disorders e.g. neoplasms of the breast or liver,  
 CC cardiovascular disorders e.g. cardiac arrest, cerebrovascular disorders  
 CC e.g. cerebral ischemia, angiodenesis, nervous system disorders e.g.  
 CC Alzheimer's disease, infections caused by bacteria, viruses and fungi and  
 CC ocular disorders e.g. corneal infection. The polypeptides can also be  
 CC used to aid wound healing and epithelial cell proliferation, to prevent  
 CC skin aging due to sunburn, to maintain organs before transplantation, for  
 CC supporting cell culture of primary tissues, to regenerate tissues and in  
 CC chemotaxis. The polypeptides can also be used as a food additive or  
 CC preservative to increase or decrease storage capabilities. The present  
 CC sequence represents a human secreted protein.  
 CC  
 XX  
 SQ Sequence 353 AA;  
 XX  
 Query Match 100.0%; Score 228; DB 22; Length 353;  
 Best Local Similarity 100.0%; Pred. No. 3.1e-222;  
 Matches 228; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 MRLTSLVVLGTAFTYVLAAGWVRLISAFLPARFOALDDRLCYQSWVLEFFENYTG 60  
 DB 1 MYLLPSVVLGTAFTYVLAAGWVRLISAFLPARFOALDDRLCYQSWVLEFFENYTG 60  
 QY 61 VQILLYGDLPKKNENITILANHOSVDWIVADILAIRONALGHVRYLKEGLKWLPLYGC 120  
 DB 61 VQILLYGDLPKKNENITILANHOSVDWIVADILAIRONALGHVRYLKEGLKWLPLYGC 120  
 QY 121 YFAOHGIIYVKRSKAKFNEKEMRNKLOSVDAGTPMVLVIFPEGTRNPEQTKVLSAQAR 180  
 DB 121 YFAOHGIIYVKRSKAKFNEKEMRNKLOSVDAGTPMVLVIFPEGTRNPEQTKVLSAQAR 180  
 QY 181 AAOAGLAVLKHLVTPRIKATHVAFDCMKNYLDAIVDTVYVEEGRDGG 228  
 DB 181 AAOAGLAVLKHLVTPRIKATHVAFDCMKNYLDAIVDTVYVEEGRDGG 228

## RESULT 5

AAB85552 standard; protein: 353 AA.

AAB85552;

25-SEP-2001 (first entry)

Human secreted protein (clone id HSL1A81).

Secreted protein: immunosuppressive; antiarthritic; antirheumatic;  
 antiapoptotic; cytoprotective; cardiact; vasotropic; cerebroprotective;  
 neurotropic; neuroprotective; antibacterial; virucide; fungicide; human;  
 ophthalmological; gene therapy.

Homo sapiens.

MO200155430-A1.

02-AUG-2001.

17-JAN-2001; 2001WO-US01431.

31-JAN-2000; 2000US-0179065.

04-FEB-2000; 2000US-0180628.

12-SEP-2000; 2000US-0231968.

(HUMA-) HUMAN GENOME SCI INC.

Rosen CA, Komatsu GA, Baker KP, Birse CE, Soppet DR, Olsen HS;

Moore PA, Wei P, Ebner R, Duan DR, Shi Y, Choi GH, Fascella M;

N-J, Ruben SM, Barash SC;

WPI: 2001-476220/51.

N-PSDB; AAH46962.

17 Isolated nucleic acid molecules encoding human secreted proteins,  
 used to preventing, treating or ameliorating a medical condition

Claim 11; Page 465-466; 482pp; English.

The invention provides novel human secreted proteins and polynucleotides  
 encoding them. The secreted proteins can be expressed by standard  
 recombinant methodology. The secreted proteins and polynucleotides are  
 used to prevent, treat or ameliorate a medical condition in e.g. humans,  
 mice, rabbits, goats, horses, cats, dogs, chickens or sheep. They can  
 also be used in diagnosing a pathological condition. The antibodies to  
 the proteins can also be used in alleviating symptoms associated with the  
 disorders and in diagnostic immunoassays e.g. radioimmunoassays or enzyme  
 linked immunosorbent assays (ELISA). Disorders which are diagnosed or  
 treated include autoimmune diseases e.g. rheumatoid arthritis,  
 hyperproliferative disorders e.g. neoplasms of the breast or liver,  
 cardiovascular disorders e.g. cardiac arrest, cerebrovascular disorders  
 e.g. cerebral ischemia, angiogenesis, nervous system disorders e.g.  
 Alzheimer's disease, infections caused by bacteria, viruses and fungi and  
 other disorders e.g. corneal infection. The polypeptides can also be  
 used to aid wound healing and epithelial cell proliferation, to prevent  
 skin aging due to sunburn, to maintain organs before transplantation, for  
 supporting cell culture of primary tissues, to regenerate tissues and in  
 chemotaxis. The polypeptides can also be used as a food additive or  
 preservative to increase or decrease storage capabilities. The present  
 sequence represents a human secreted protein.

Sequence 353 AA;

Query Match 100.0%; Score 228; DB 22; Length 353;

Best Local Similarity 100.0%; Pred. No. 3.1e-222;

Matches 228; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 MRYLLPSVLLGTAPTYVLAAGWVRLLSAPLPARFYQALDRLCYQSWLFFFEYNTG 60  
 1 MYLLPSVLLGTAPTYVLAAGWVRLLSAPLPARFYQALDRLCYQSWLFFFEYNTG 60

QY 61 VQILLYGDLPKKNKEIIYLANHGSTVDWIVADILAIROMALGHVRYLKEGLKMLPYGC 120

DB 61 VQILLYGDLPKKNKEIIYLANHGSTVDWIVADILAIQNALGHVRYLKEGLKMLPYGC 120

QY 121 YFAOHGIGYVRSKAFNFKEMRNKLOSYVDAGTGMVLYIFPEGRYRNEPQRYVLSAQAF 180

DB 121 YFAOHGIGYVRSKAFNFKEMRNKLOSYVDAGTGMVLYIFPEGRYRNEPQRYVLSAQAF 180

QY 181 AAGRLAVLKHVLPRIKATVAFDCMKKNYDAIYDTVYVEGKDDG 228

DB 181 AAGRLAVLKHVLPRIKATVAFDCMKKNYDAIYDTVYVEGKDDG 228

## RESULT 6

AAV36729 standard; Protein: 364 AA.

AAV36729;

27-SEP-1999 (first entry)

Human PGI protein sequence.

PGI gene; biallelic marker; human; PSA; PGI-related biallelic marker;

cancer; prostate cancer; diagnosis; therapy; prostate specific antigen.

Homo sapiens.

W09932644-A2.

01-JUL-1999.

22-DEC-1998; 98WO-IB02133.

09-SEP-1998; 98US-0099658.

22-DEC-1997; 97US-0996306.

(GEST) GENSET.

Blumenfeld M, Bougueleret L, Chumakov I, Cohen D;

WPI: 1999-405178/34.

N-PSDB; AA200870.

Claim 7; Page 190-191; 365pp; English.

The invention relates to a mammalian PGI gene and protein, and a set of  
 PGI biallelic markers. The PGI polynucleotide and biallelic markers are  
 used in a hybridisation assay, a sequencing assay, or in an  
 allele-specific amplification assay for determining the identity of a  
 nucleotide at a PGI-related biallelic marker. The methods can be used to  
 detect and to assess the risk of developing cancer or prostate cancer.  
 Early-stage diagnosis of prostate cancer relies on prostate specific  
 antigen (PSA) dosage. However, the effectiveness of this is limited due  
 to its inability to discriminate between malignant and non-malignant  
 affections of the organ. A need exists for both a reliable diagnostic  
 procedure which would enable early-stage diagnosis, and for preventative  
 and curative treatments of the disease. The PGI gene can be used for  
 detection of prostate cancer, and the risk of developing it in the  
 future, and can also be used to determine therapies for the disease.

Sequence 364 AA;

Query Match 100.0%; Score 228; DB 20; Length 364;

Best Local Similarity 100.0%; Pred. No. 3.2e-222;

Matches 228; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 MRYLLPSVLLGTAPTYVLAAGWVRLLSAPLPARFYQALDRLCYQSWLFFFEYNTG 60

[illegible]

FT	Modified-site	130..135	
FT	/note= "N-myristoylated"		
FT	Modified-site	330..335	
FT	/note= "N-myristoylated"		
FT	Modified-site	352..357	
FT	/note= "N-myristoylated"		
FT	Modified-site	230..343	
FT	/note= "Amlidation"		
FT	Peptide	105..126	
FT	/note= "predicted leucine zipper"		
XX			
PN	W0200173051-A2.		
XX			
PD	04-OCT-2001.		
XX			
PF	26-MAR-2001; 2001WO-US09633.		
XX			
PR	24-MAR-2000; 2000US-192092P.		
XX			
PA	(MILL-) MILLENNIUM PHARM INC.		
XX			
PI	Meyers RA, Rudolph-Owen L, Macbeth KJ, Tsai F, Williamson M;		
XX			
DR	WPI: 2001-626267/72.		
XX			
DR	N-PSDB: AAH27094.		
XX			
PT	New polynucleotide encodes a polypeptide for the diagnosis and		
PT	treatment of cancer and to alter cellular lipid metabolism comprises a		
PT	polynucleotide encoding human acyltransferases designated 46743 and		
PT	27417		
XX			
PS	Claim 9; Fig 8A-B; 150pp; English.		
XX			
CC	The present sequence is that of a novel human protein, termed		
CC	27417, which shows the structural characteristics of members of the		
CC	acyltransferase family, including an acyltransferase domain. The		
CC	invention provides novel acyltransferase 27417 nucleic acids and		
CC	polypeptides, as well as methods for detecting their presence, and		
CC	methods for screening for compounds that modulate their expression		
CC	or activity. Such compounds can be used to treat conditions		
CC	related to aberrant activity or expression of 27417 protein or		
CC	nucleic acid, such as conditions involving deficient cellular		
CC	proliferation, migration and/or differentiation. Also provided are		
CC	methods for inhibiting the proliferation or migration, or inducing		
CC	the killing, of a 27417-expressing cell, e.g. a hyperproliferative		
CC	and/or metastatic cell, by contact with a compound that modulates		
CC	the activity or expression of the 27417 protein or nucleic acid.		
CC	The 27417-expressing cell is found in the healthy or diseased		
CC	heart, blood vessels, kidney, skeletal muscle, brain or liver, or		
CC	especially in a solid tumour, a soft tissue tumour or a metastatic		
CC	lesion, a sarcoma, a carcinoma or an adenocarcinoma, and in		
CC	particular in a hyperproliferative and/or metastatic cell found in		
CC	ovarian, brain, colon or lung cancer. The compounds may also be		
CC	used to modulate lipid metabolism in a 27417-expressing cell.		
CC	Methods of disease diagnosis, e.g. by determining the		
CC	presence of a genetic alteration in a 27417 polypeptide, and for		
CC	evaluating the efficacy of a treatment of a disorder, are also		
CC	provided.		
XX			
XX			
SQ	Sequence 364 AA:		
	Query Match 100.0%; Score 228; DB 22; Length 364;		
	Best Local Similarity 100.0%; Pred. NO. 3.2e-222;		
	Matches 228; Conservative 0; Mismatches 0; Indels 0; Gaps		
OY	1 MRYLLPSVLLGTATFYVLAWGVWRLLSAFLPARYQALDDRLKXCVCYOSMVLFFENYTG 60		
DB	12 mryllpsvlllgtatptrylawnvwlslsaflparfygaldrllycvysgmvlffennytg 71		
OY	61 VOILLYGLPKKNNIIVLANHGSTVDWIVADIALIRONALGHRVRLKEGLKMLPLYGC 120		
DB	72 voilllygdlpkokennilylanhgstvdwlyvadialirgnalghrvyrlkeglkmlplygc 131		

Query	Match	Score	DB	Length
1	KRYLLPVSLLCTAPTYLVAMGWRLLSAFLPARFYQALDDRLCYVQSMVLEFFENYTG	100.0%	100.0%	364
2	Best Local Similarity	100.0%	Pred. No. 3,2e-222	
3	Matches	228	Conservative	0; Mismatches 0; Indels 0; Gaps 0

Query Match	Best Local Similarity	Score	DB	Length
Matches 185; Conservative	100.0%; Pred. No. 5.6e-179; Mismatches 0; Indels 0; Gaps 0;	81.1%;	DB 20;	Length 185;

OY 1 MRYLLPSVLLGTPTVYLAIVGVRLLISAFIPARFYQALDDRLXCYQSVNLEFFENYTG 60  
 Db 1 mryllpsvllgtptvylawgvrrllsafirpariyqalddrlxcvqsmvllffenytg 60  
 OY 61 VOIILLYGDLPEKNKNNIITYLANHSTYDMIVADIATIRQNALGHVRYLYKEGLKMLPLYGC 120  
 Db 61 vqililglydlpknknniilylanhsgtvmivadiatirqnalghvrylyksglkmlplygc 120  
 OY 121 YFAOHGIGYVYKRSKAFNEKEKRNKLQSYVDAGTIPMYLVIFPEGTRNPEQTKVLSAQAF 180  
 Db 121 yfaohgiglyvyrskafnekekrmnklqsyvdagtpmylvifpegtrynpeqtkvlsasqaf 180  
 OY 181 AAORG 185  
 Db 181 aagrg 185  
 RESULT 10  
 AAY36751  
 ID AAY36751 standard; Protein: 315 AA.  
 XX AAY36751;  
 AC  
 XX  
 XX 27-SEP-1999 (first entry)  
 DT  
 XX  
 DE PGI splice variant allele protein sequence.  
 XX  
 KW PGI gene; biallelic marker; human; PSA; PGI-related biallelic marker;  
 XX cancer; prostate cancer; diagnosis; therapy; prostate specific antigen.  
 XX  
 OS Homo sapiens.  
 OS  
 PN W09932644-A2.  
 PN  
 PD 01-JUL-1999.  
 PD  
 PF 22-DEC-1998; 98WO-IB02133.  
 PF  
 XX 09-SEP-1998; 980US-00996658.  
 PR 22-DEC-1997; 97US-0096306.  
 XX  
 PA (GEST ) GENSET.  
 XX  
 XX Blumenfeld M, Bougueleret L, Chumakov I, Cohen D;  
 PI  
 DR WPI: 1999-405178/34.  
 DR N-PSDB; AAZ00977.  
 XX  
 PT Use of a prostate cancer associated gene and biallelic markers  
 PT derived from it  
 PS  
 PS Claim 7; Page 253-254; 385pp; English.  
 CC  
 CC The invention relates to a mammalian PGI gene and protein, and a set of  
 CC PGI biallelic markers. The PGI polynucleotide and biallelic markers are  
 CC used in a hybridisation assay, a sequencing assay, or in an  
 CC allele-specific amplification assay for determining the identity of a  
 CC nucleotide at a PGI-related biallelic marker. The methods can be used to  
 CC detect and to assess the risk of developing cancer or prostate cancer.  
 CC Early-stage diagnosis of prostate cancer relies on prostate specific  
 CC antigen (PSA) dosage. However, the effectiveness of this is limited due  
 CC to its inability to discriminate between malignant and non-malignant  
 CC affections of the organ. A need exists for both a reliable diagnostic  
 CC procedure which would enable early-stage diagnosis, and for preventative  
 CC and curative treatments of the disease. The PGI gene can be used for  
 CC detection of prostate cancer, and the risk of developing it in the  
 CC future, and can also be used to determine therapies for the disease.  
 CC  
 Sequence 315 AA;  
 90

Query Match

81.18; Score 185; DB 20; Length 315;

Best Local Similarity 100.0%; Pred. No. 9,2e-179;  
Matches 185; Conservative 0; Mismatches 0; Indels 0; Gaps 0.

Qy	1	MRYLPSVLLGTPPTVYLAQVWRLLSAFLPAEFYOALDDRLCYCQVSNVLEPFEENYTG	60
Db	1	mcylpssvlllgtpapcyylawgwrlsalfparfygaldarlycyyqsmvlfefenytg	60
Qy	61	VOILLGDLPRKNENIYVLANHNSVTVMIVADILAIQNALGHNRYLAKGSLKMLPLYGC	120
Db	61	vqlllygdlprknkniylulanhgstvmvadvdilaifqnalghnryylkqgllwplpygc	120
Qy	121	YFAOHGIGIVKRSAKFNENKEMRNKLQSTVDAGTPTMYLVIFPEEGRTNVPEDTKVLSAQAF	180
Db	121	yfaqhgiglyvkrsakfnekemrnklqsgyvdagtpmylvilfpegtynpqrctkylssaqaf	180
Qy	181	AAQRG 185	
Db	181	aqrq 185	

## RESULT 11

ID AAY36752 standard; Protein; 300 AA.

AC AAY36752;

DT 27-SEP-1999 (first entry)

DE PGI splice variant allele protein sequence.

PGI gene; biallelic marker; human; PSA; PGI-related biallelic marker; cancer; prostate cancer; diagnosis; therapy; prostate specific antigen

OS Homo sapiens.

PN W09932644-A2.

PD 01-JUL-1999.

PF 22-DEC-1998; 98WO-IB02133.

PR 09-SEP-1998; 98US-0099658.

XX

XX

XX

DR N-PSDB; AAZ00978.

PT use of a prostate cancer associated gene and biallelic markers

XX

The invention relates to a mammalian PGI gene and protein, and a set of PGI biallelic markers. The PGI polynucleotide and biallelic markers are used in a hybridisation assay, a sequencing assay, or in an allele-specific amplification assay for determining the identity of a nucleotide at a PGI-related biallelic marker. The methods can be used to detect and to assess the risk of developing cancer or prostate cancer. Early-stage diagnosis of prostate cancer relies on prostate specific antigen (PSA) dosage. However, the effectiveness of this is limited due to its inability to discriminate between malignant and non-malignant affections of the organ. A need exists for both a reliable diagnostic procedure which would enable early-stage diagnosis, and for preventative and curative treatments of the disease. The PGI gene can be used for detection of prostate cancer, and the risk of developing it in the future, and can also be used to determine therapies for the disease.

Sequence 300 AA;

**James A. McHugh**



PR	14-AUG-2000	2000US-0225271
PR	14-AUG-2000	2000US-0225447
PR	14-AUG-2000	2000US-0225757
PR	14-AUG-2000	2000US-0225758
PR	14-AUG-2000	2000US-0225759
PR	18-AUG-2000	2000US-0226279
PR	22-AUG-2000	2000US-0226621
PR	22-AUG-2000	2000US-0226868
PR	22-AUG-2000	2000US-0227182
PR	22-AUG-2000	2000US-0227100
PR	30-AUG-2000	2000US-0228982
PR	01-SEP-2000	2000US-0229287
PR	01-SEP-2000	2000US-0229343
PR	01-SEP-2000	2000US-0229344
PR	01-SEP-2000	2000US-0229345
PR	05-SEP-2000	2000US-0229509
PR	05-SEP-2000	2000US-0229513
PR	06-SEP-2000	2000US-0230437
PR	06-SEP-2000	2000US-0230348
PR	08-SEP-2000	2000US-0231242
PR	08-SEP-2000	2000US-0231243
PR	08-SEP-2000	2000US-0231244
PR	08-SEP-2000	2000US-0231413
PR	08-SEP-2000	2000US-0231414
PR	08-SEP-2000	2000US-02312080
PR	12-SEP-2000	2000US-0231068
PR	12-SEP-2000	2000US-0231968
PR	14-SEP-2000	2000US-0232397
PR	14-SEP-2000	2000US-0232398
PR	14-SEP-2000	2000US-0232399
PR	14-SEP-2000	2000US-0232399
PR	14-SEP-2000	2000US-0232400
PR	14-SEP-2000	2000US-0232401
PR	14-SEP-2000	2000US-0233063
PR	14-SEP-2000	2000US-0233064
PR	14-SEP-2000	2000US-0233065
PR	21-SEP-2000	2000US-0234223
PR	21-SEP-2000	2000US-0234274
PR	25-SEP-2000	2000US-0234998
PR	25-SEP-2000	2000US-0234998
PR	26-SEP-2000	2000US-0235484
PR	27-SEP-2000	2000US-0235834
PR	29-SEP-2000	2000US-0236327
PR	29-SEP-2000	2000US-0236327
PR	29-SEP-2000	2000US-0236359
PR	29-SEP-2000	2000US-0236368
PR	29-SEP-2000	2000US-0236370
PR	02-OCT-2000	2000US-0236802
PR	02-OCT-2000	2000US-0237037
PR	02-OCT-2000	2000US-0241961
PR	02-OCT-2000	2000US-0237038
PR	02-OCT-2000	2000US-0237039
PR	02-OCT-2000	2000US-0237040
PR	13-OCT-2000	2000US-0239935
PR	13-OCT-2000	2000US-0239937
PR	20-OCT-2000	2000US-0240260
PR	20-OCT-2000	2000US-0241921
PR	20-OCT-2000	2000US-0241785
PR	20-OCT-2000	2000US-0241865
PR	20-OCT-2000	2000US-0241787
PR	20-OCT-2000	2000US-0241808
PR	20-OCT-2000	2000US-0241809
PR	20-OCT-2000	2000US-0241816
PR	01-NOV-2000	2000US-0244517
PR	08-NOV-2000	2000US-0244674
PR	08-NOV-2000	2000US-0244675
PR	08-NOV-2000	2000US-0244676
PR	08-NOV-2000	2000US-0246476
PR	08-NOV-2000	2000US-0246477
PR	08-NOV-2000	2000US-0246478
PR	08-NOV-2000	2000US-0246523
PR	08-NOV-2000	2000US-0246524
PR	08-NOV-2000	2000US-0246525
PR	08-NOV-2000	2000US-0246526
PR	08-NOV-2000	2000US-0246527
PR	08-NOV-2000	2000US-0246528

08-NOV-2000; 2000US-0246528.  
PR 08-NOV-2000; 2000US-0246532.  
PR 08-NOV-2000; 2000US-0246539.  
PR 08-NOV-2000; 2000US-0246609.  
PR 08-NOV-2000; 2000US-0246610.  
PR 08-NOV-2000; 2000US-0246611.  
PR 08-NOV-2000; 2000US-0246613.  
PR 17-NOV-2000; 2000US-0249207.  
PR 17-NOV-2000; 2000US-0249208.  
PR 17-NOV-2000; 2000US-0249209.  
PR 17-NOV-2000; 2000US-0249210.  
PR 17-NOV-2000; 2000US-0249211.  
PR 17-NOV-2000; 2000US-0249212.  
PR 17-NOV-2000; 2000US-0249213.  
PR 17-NOV-2000; 2000US-0249214.  
PR 17-NOV-2000; 2000US-0249215.  
PR 17-NOV-2000; 2000US-0249216.  
PR 17-NOV-2000; 2000US-0249217.  
PR 17-NOV-2000; 2000US-0249218.  
PR 17-NOV-2000; 2000US-0249244.  
PR 17-NOV-2000; 2000US-0249245.  
PR 17-NOV-2000; 2000US-0249264.  
PR 17-NOV-2000; 2000US-0249265.  
PR 17-NOV-2000; 2000US-0249297.  
PR 17-NOV-2000; 2000US-0249299.  
PR 17-NOV-2000; 2000US-0249300.  
PR 01-DEC-2000; 2000US-0250160.  
PR 01-DEC-2000; 2000US-0250391.  
PR 05-DEC-2000; 2000US-0251030.  
PR 05-DEC-2000; 2000US-0251988.  
PR 05-DEC-2000; 2000US-0256719.  
PR 06-DEC-2000; 2000US-0251479.  
PR 08-DEC-2000; 2000US-0251856.  
PR 08-DEC-2000; 2000US-0251868.  
PR 08-DEC-2000; 2000US-0251869.  
PR 08-DEC-2000; 2000US-0251989.  
PR 08-DEC-2000; 2000US-0251990.  
PR 11-DEC-2000; 2000US-0254097.  
PR 05-JAN-2001; 2001US-0259678.  
XX  
XX  
PA (HUMA-) HUMAN GENOME SCI INC.  
XX  
XX  
PI Rosen CA, Barash SC, Ruben SM,  
XX  
XX WPI; 2001-457716/49.  
DR N-PSDB; AAD16620.  
XX  
XX  
PT Isolated lipid metabolism polypeptide for screening to identify  
PT antagonists and agonists that may enhance or block activities mediated  
PT by lipid metabolism proteins and also for testing and detection e.g.  
PI diagnosis -  
XX  
PS Claim 11; SEQ ID NO 47; 453pp; English.  
XX  
XX AAD16606-AAD16627 represent cDNAs corresponding to human lipid metabolism  
CC protein genes, and AAE09639-AAE09660 represent the proteins they encode.  
CC AAD16628-AAD16647 represent human lipid metabolism genomic DNAs. The  
CC genes and their corresponding proteins are useful for preventing,  
CC treating or ameliorating medical conditions such as immune disorders,  
CC inflammatory disorders (e.g., conjunctivitis, glossitis), autoimmune  
CC diseases (e.g., rheumatoid arthritis, systemic lupus erythematosus),  
CC respiratory disorders (asthma, allergy), CNS disorders (e.g., Alzheimer's  
CC disease, Parkinson's disease), haematopoietic disorders, cardiovascular  
CC disorders (e.g., atherosclerosis, coronary disease), hyperlipidaemias,  
CC hypolipidaemia, lipidoses, familial hypercholesterolaemia, Gaucher's  
CC disease, Tay-sach's disease, diabetes mellitus, cancer and other enzyme  
CC abnormalities. They are used for the prevention of mental retardation  
CC and the enlargement of spleen and liver. They are used in gene therapy  
CC and antisense therapy. The present sequence represents a human lipid  
CC metabolism protein of the invention.  
CC Note: The sequence data for this patent did not form part of the printed  
CC specification, but was obtained in electronic format directly from WIPO  
CC at ftp.wipo.int/pub/published\_pct\_sequences.  
XX





PT Novel nucleic acids and polypeptides, useful for treating disorders  
PT such as central nervous system injuries -

XX  
PS Example 2: SEQ ID NO 6308; 10078bp; English.

CC The invention relates to human nucleic acids (AA157798-AA161369) and  
CC the encoded polypeptides (AAM38642-AAM42213) with nootropic,  
CC immunosuppressant and cytostatic activity. The polynucleotides are useful  
CC in gene therapy. A composition containing a polypeptide or polynucleotide  
CC of the invention may be used to treat diseases of the peripheral nervous  
CC system, such as peripheral nervous injuries, peripheral neuropathy and  
CC localised neuropathies and central nervous system diseases, such as  
CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic  
CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the  
CC utilisation of the activities such as: Immune system suppression,  
CC Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic  
CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,  
CC assays for receptor activity, arthritis and inflammation, leukaemias and  
CC C.N.S disorders.  
CC Note: The sequence data for this patent did not form part of the printed  
CC specification.

XX  
SQ Sequence 372 AA:

Query Match 55.7%; Score 127; DB 22; Length 372;

Best Local Similarity 99.6%; Pred No. 5.2e-120;

Matches 227; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 MRYLLPSVVLGTAFTYVLANGVWRLLSAFLPARFYQALDRLCYVQSMVLEFFENYTG 60  
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
DB 20 MYLLPSVVLGTAFTYVLANGVWRLLSAFLPARFYQALDRLCYVQSMVLEFFENYTG 79  
OY 61 VQILLXGDLPRKKNENIYIANQSTYVDWIVADILAIRONALGHVRYVLEKGLKMLPLCYC 120  
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
DB 80 VQILLXGDLPRKKNENIYIANQSTYVDWIVADILAIRONALGHVRYVLEKGLKMLPLCYC 139  
OY 121 YFAQHGAGIVKRSKAKFNEKEMRNKLOSIVDAGCPMYLVTFPEGSTRYNPEOTKVLSSQAF 180  
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
DB 140 YFAQHGAGIVKRSKAKFNEKEMRNKLOSIVDAGCPMYLVTFPEGSTRYNPEOTKVLSSQAF 199  
OY 181 AAQRGLAVLKHLVLPRIKATHVAFDCKMKNYLDIYDVTVVYEGKDDGG 228  
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
DB 200 AAGRGLAVLKHLVLPRIKATHVAFDCKMKNYLDIYDVTVVYEGKDDGG 247

Search completed: August 28, 2002, 11:18:45  
Job time: 583 sec

**THIS PAGE BLANK (USPTO)**

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 28, 2002, 11:19:21 ; Search time 30.82 Seconds  
(without alignments)  
180.696 Million cell updates/sec

Title: US-09-853-526-70

Perfect score: 228

Sequence: 1 MRYLLPSVILGTAPTYVLA.....NYLDAIVDVTVYEGKDDG 228

Scoring table:

Gapop 60.0 , Gapext 60.0

Searched: 231628 seqs, 24425594 residues

Word size : 0

Total number of hits satisfying chosen parameters: 231628

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database :

- 1: /cgn2\_6/ptodata/2/1aa/5A.COMB.pep:\*
- 2: /cgn2\_6/ptodata/2/1aa/5B.COMB.pep:\*
- 3: /cgn2\_6/ptodata/2/1aa/6A.COMB.pep:\*
- 4: /cgn2\_6/ptodata/2/1aa/6B.COMB.pep:\*
- 5: /cgn2\_6/ptodata/2/1aa/PTCUTS.COMB.pep:\*
- 6: /cgn2\_6/ptodata/2/1aa/Dackfile1.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	228	100.0	228	4	US-09-338-907-70
2	228	100.0	228	4	US-09-218-207-70
3	228	100.0	353	2	US-08-996-306-4
4	228	100.0	353	4	US-09-338-907-4
5	228	100.0	353	4	US-09-218-207-4
6	228	100.0	364	2	US-08-996-306-5
7	228	100.0	364	4	US-09-338-907-5
8	228	100.0	364	4	US-09-218-207-5
9	185	81.1	185	4	US-09-338-907-136
10	185	81.1	185	4	US-09-218-207-136
11	185	81.1	315	4	US-09-338-907-134
12	185	81.1	315	4	US-09-218-207-134
13	184	80.7	300	4	US-09-338-907-135
14	184	80.7	300	4	US-09-218-207-135
15	154	67.5	182	4	US-09-338-907-133
16	154	67.5	182	4	US-09-218-207-133
17	105	46.1	291	4	US-09-338-907-127
18	105	46.1	291	4	US-09-218-207-127
19	85	37.3	97	4	US-09-338-907-132
20	85	37.3	97	4	US-09-218-207-132
21	74	32.5	261	4	US-09-338-907-128
22	74	32.5	261	4	US-09-218-207-128
23	62	27.2	66	4	US-09-338-907-131
24	62	27.2	66	4	US-09-218-207-131
25	62	27.2	68	4	US-09-338-907-130
26	62	27.2	68	4	US-09-218-207-130
27	62	27.2	77	4	US-09-338-907-125

28	62	27.2	77	4	US-09-218-207-125	Sequence 125, App
29	62	27.2	90	4	US-09-338-907-129	Sequence 129, App
30	62	27.2	90	4	US-09-218-207-129	Sequence 129, App
31	62	27.2	238	4	US-09-338-907-126	Sequence 126, App
32	62	27.2	238	4	US-09-338-907-126	Sequence 126, App
33	62	27.2	238	4	US-09-338-907-126	Sequence 126, App
34	32	14.0	354	4	US-09-218-207-74	Sequence 74, Appl
35	9	3.9	9	2	US-08-996-306-15	Sequence 15, Appl
36	9	3.9	9	4	US-09-338-907-15	Sequence 15, Appl
37	9	3.9	9	4	US-09-218-207-15	Sequence 15, Appl
38	8	3.5	9	2	US-08-996-306-18	Sequence 18, Appl
39	8	3.5	9	4	US-09-338-907-18	Sequence 18, Appl
40	8	3.5	9	4	US-09-218-207-18	Sequence 18, Appl
41	7	3.1	12	2	US-08-224-625-1	Sequence 1, Appl1
42	7	3.1	12	2	US-08-254-404-1	Sequence 1, Appl1
43	7	3.1	12	2	US-08-327-451E-1	Sequence 1, Appl1
44	7	3.1	12	2	US-08-458-109-1	Sequence 1, Appl1
45	7	3.1	12	3	US-08-231-196-1	Sequence 1, Appl1

ALIGNMENTS

RESULT 1  
US-09-338-907-70  
; Sequence 70, Application US/09338907  
; Patent No. 6265546  
; GENERAL INFORMATION:  
; APPLICANT: Cohen, Daniel  
; APPLICANT: Blumenfeld, Marla  
; APPLICANT: Ilya, Chumakov  
; APPLICANT: Bougueleret, Lydie  
; TITLE OF INVENTION: PROSTATE CANCER GENE  
; FILE REFERENCE: GENSET.18CP1CP  
; CURRENT APPLICATION NUMBER: US/09/338, 907  
; EARLIER FILING DATE: 1999-06-23  
; EARLIER APPLICATION NUMBER: 08/996, 306  
; EARLIER FILING DATE: 1997-12-22  
; EARLIER APPLICATION NUMBER: 60/099, 658  
; EARLIER FILING DATE: 1998-09-09  
; EARLIER APPLICATION NUMBER: 09/218, 207  
; EARLIER FILING DATE: 1998-12-22  
; NUMBER OF SEQ ID NOS: 578  
; SOFTWARE: Patent.pm  
; SEQ ID NO 70  
; LENGTH: 228  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-338-907-70

Query Match 100.0%; Score 228; DB 4; Length 228;  
Best Local Similarity 100.0%; Pred. No. 1.1e-221; Indels 0; Gaps 0;  
Matches 228; Conservative 0; Mismatches 0;

QY	1	MRYLLPSVILGTAPTYVLA	WGRLLSAFLPAFQYALDDRLCYQSWLFFFEWYTG	60
Db	1	MRYLLPSVILGTAPTYVLA	WGRLLSAFLPAFQALDDRLCYQSWLFFFEWYTG	60
QY	61	VQILLXGDLPRKNNITIIYLANHSTVDWIVADLITRQNALGHVRYLKLKMLPLYGC	120	
Db	61	VQILLXGDLPRKNNITIIYLANHSTVDWIVADLITRQNALGHVRYLKLKMLPLYGC	120	
QY	121	YFAOHGIIYKRSKAKFNEKMRKKLOSVDAGTIPMYLIFPESTRINPEOTKYLSSOAF	180	
Db	121	YFAOHGIIYKRSKAKFNEKMRKKLOSVDAGTIPMYLIFPESTRINPEOTKYLSSOAF	180	
QY	181	AAORGLAVLKHLVLPRIKATHVAFDCCKNYLDAIVDTVYVEGKDDG	228	
Db	181	AAORGLAVLKHLVLPRIKATHVAFDCCKNYLDAIVDTVYVEGKDDG	228	

RESULT 2

US-09-218-207-70  
; Sequence 70, Application US/09218207  
; Patent No. 6346381  
; GENERAL INFORMATION:  
; APPLICANT: Cohen, Daniel  
; APPLICANT: Blumenfeld, Marta  
; APPLICANT: Ilya, Chumakov  
; APPLICANT: Bougueleret, Lydie  
; TITLE OF INVENTION: Prostate cancer gene  
; FILE REFERENCE: GENSET.018CPI  
; CURRENT APPLICATION NUMBER: US/09/218, 207  
; EARLIER FILING DATE: 1993-12-22  
; EARLIER APPLICATION NUMBER: 08/996, 306  
; EARLIER FILING DATE: 1997-12-22  
; EARLIER APPLICATION NUMBER: 60/099, 658  
; EARLIER FILING DATE: 1996-09-09  
; NUMBER OF SEQ ID NOS: 578  
; SOFTWARE: Patent.pm  
; SEQ ID NO 70  
; LENGTH: 228  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-218-207-70

Query Match 100.0%; Score 228; DB 4; Length 228;  
Best Local Similarity 100.0%; Pred. No. 1.1e-221;  
Matches 228; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MRYLLPSVLLGTAPTYVLLANGVRLLSAFLPARFYQALDRLKCYVQSMVLPFFENTG 60  
DB 1 MRYLLPSVLLGTAPTYVLLANGVRLLSAFLPARFYQALDRLKCYVQSMVLPFFENTG 60  
QY 61 VQILLYGDLPRKNENIYILNHOSTVDIVADILAIRONALGHRYVYKEGLKMLPL/GC 120  
DB 61 VQILLYGDLPRKNENIYILNHOSTVDIVADILAIRONALGHRYVYKEGLKMLPL/GC 120  
QY 121 YFAOHGCIYVRSKAFENKERNKLSQSYVDAGTPMYLVIFPEGTRYNDEQTRVLSASQAF 180  
DB 121 YFAOHGCIYVRSKAFENKERNKLSQSYVDAGTPMYLVIFPEGTRYNDEQTRVLSASQAF 180  
QY 181 AAOGLAVLKHVLPRIKATIVAFDCKMKNYLDATYDVTVYVEGKDDGG 228  
DB 181 AAOGLAVLKHVLPRIKATIVAFDCKMKNYLDATYDVTVYVEGKDDGG 228

RESULT 3  
US-08-996-306-4  
; Sequence 4, Application US/08996306  
; Patent No. 5945522  
; GENERAL INFORMATION:  
; APPLICANT: Cohen, Daniel  
; APPLICANT: Chumakov, Ilya  
; APPLICANT: Blumenfeld, Marta  
; APPLICANT: Bougueleret, Lydie  
; TITLE OF INVENTION: Prostate cancer gene  
; NUMBER OF SEQUENCES: 68  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Knobbe, Martens, Olson & Bear  
; STREET: 501 West Broadway  
; CITY: San Diego  
; STATE: California  
; COUNTRY: USA  
; ZIP: 92101-3505  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy Disk  
; OPERATING SYSTEM: IBM PC compatible  
; SOFTWARE: Word  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/996, 306  
; FILING DATE:  
; CLASSIFICATION:

ATTORNEY/AGENT INFORMATION:  
NAME: Israelsen, Ned A.  
REGISTRATION NUMBER: 29,655  
REFERENCE/DOCKET NUMBER: GENSET.018A  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (619) 235-8550  
TELEFAX: (619) 235-0176  
INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 353 amino acids  
TYPE: AMINO ACID  
STRANDEDNESS: SINGLE  
TOPOLOGY: LINEAR  
MOLECULE TYPE: PROTEIN  
ORIGINAL SOURCE:  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: potential Transmembrane helix  
LOCATION: 1..33  
IDENTIFICATION METHOD: Rao and Argos method  
FEATURE:  
NAME/KEY: potential Transmembrane helix  
LOCATION: 4..20  
IDENTIFICATION METHOD: Klein, Kanehisa and Delisi method  
FEATURE:  
NAME/KEY: potential Transmembrane helix  
LOCATION: 4..24  
IDENTIFICATION METHOD: Eisenberg, Schwarz, Komarony and Wall method  
FEATURE:  
NAME/KEY: potential N-myristoylation site  
LOCATION: 12  
IDENTIFICATION METHOD: prosite match  
FEATURE:  
NAME/KEY: potential Transmembrane helix  
LOCATION: 50..70  
IDENTIFICATION METHOD: Eisenberg, Schwarz, Komarony and Wall method  
FEATURE:  
NAME/KEY: potential N-glycosylation site  
LOCATION: 57  
IDENTIFICATION METHOD: prosite match  
FEATURE:  
NAME/KEY: potential Transmembrane helix  
LOCATION: 76..96  
IDENTIFICATION METHOD: Eisenberg, Schwarz, Komarony and Wall method  
FEATURE:  
NAME/KEY: potential Tyrosine kinase phosphorylation site  
LOCATION: 78  
IDENTIFICATION METHOD: prosite match  
FEATURE:  
NAME/KEY: potential Casein kinase II phosphorylation site  
LOCATION: 84  
IDENTIFICATION METHOD: prosite match  
FEATURE:  
NAME/KEY: potential Leucine zipper pattern  
LOCATION: 94..115  
IDENTIFICATION METHOD: prosite match  
FEATURE:  
NAME/KEY: potential N-myristoylation site  
LOCATION: 119  
IDENTIFICATION METHOD: prosite match  
FEATURE:  
NAME/KEY: potential Protein kinase C phosphorylation site  
LOCATION: 133  
IDENTIFICATION METHOD: prosite match  
FEATURE:  
NAME/KEY: potential Casein kinase II phosphorylation site  
LOCATION: 147  
IDENTIFICATION METHOD: prosite match  
FEATURE:  
NAME/KEY: potential Protein kinase C phosphorylation site  
LOCATION: 194  
IDENTIFICATION METHOD: prosite match

FEATURE:  
NAME/KEY: potential Tyrosine kinase phosphorylation site  
LOCATION: 215  
IDENTIFICATION METHOD: prosite match  
FEATURE:  
NAME/KEY: potential Tyrosine sulfatation site  
LOCATION: 221  
IDENTIFICATION METHOD: prosite match  
FEATURE:  
NAME/KEY: potential cAMP- and cGMP-dependent protein kinase phosphorylation site  
LOCATION: 223  
IDENTIFICATION METHOD: prosite match  
FEATURE:  
NAME/KEY: potential Casein kinase II phosphorylation site  
LOCATION: 225  
IDENTIFICATION METHOD: prosite match  
FEATURE:  
NAME/KEY: potential Protein kinase C phosphorylation site  
LOCATION: 306  
IDENTIFICATION METHOD: prosite match  
FEATURE:  
NAME/KEY: potential Transmembrane helix  
LOCATION: 310..330  
IDENTIFICATION METHOD: Eisenberg, Schwarz, Komarony and Wall method  
FEATURE:  
NAME/KEY: potential N-myristoylation site  
LOCATION: 319  
IDENTIFICATION METHOD: prosite match  
FEATURE:  
NAME/KEY: potential N-myristoylation site  
LOCATION: 323  
IDENTIFICATION METHOD: prosite match  
FEATURE:  
NAME/KEY: potential Amidation site  
LOCATION: 329  
IDENTIFICATION METHOD: prosite match  
FEATURE:  
NAME/KEY: potential Transmembrane helix  
LOCATION: 333..353  
IDENTIFICATION METHOD: Eisenberg, Schwarz, Komarony and Wall method  
FEATURE:  
NAME/KEY: potential N-myristoylation site  
LOCATION: 341  
IDENTIFICATION METHOD: prosite match  
FEATURE:  
NAME/KEY: potential Protein kinase C phosphorylation site  
LOCATION: 350  
IDENTIFICATION METHOD: prosite match  
US-08-996-306-4

Query Match 100.0%; Score 228; DB 2; Length 353;  
Best Local Similarity 100.0%; Pred. No. 1.6e-221;  
Matches 228; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 MRLLEPSVLLGAPPTVYVAMGWRLISAFLEPARFYQALDDRLCYVQSQWVLEFFENYTG 60  
DB 1 MRLLEPSVLLGAPPTVYVAMGWRLISAFLEPARFYQALDDRLCYVQSQWVLEFFENYTG 60  
OY 61 VOILLYGDLPKKNENITVLANHOSVDMVADILAIKRONALGHVRYVLEKGLMPLLYGC 120  
DB 61 VOILLYGDLPKKNENITVLANHOSVDMVADILAIKRONALGHVRYVLEKGLMPLLYGC 120  
OY 121 YFAOHGIGYVKSAREKEMRNKLOSVDAGTPMTLVIFPEGTIRNPEQTKVLSAQAF 180  
DB 121 YFAOHGIGYVKSAREKEMRNKLOSVDAGTPMTLVIFPEGTIRNPEQTKVLSAQAF 180  
OY 181 AAOGLAVLKHVLTPIKATHVAFDCMKNTLDAIVDTVYVEEKDGG 228  
DB 181 AAOGLAVLKHVLTPIKATHVAFDCMKNTLDAIVDTVYVEEKDGG 228

RESULT 4

US-09-338-907-4  
Sequence 4, Application US/09338907  
Patent No. 6265546  
GENERAL INFORMATION:  
APPLICANT: Cohen, Daniel  
APPLICANT: Blumenfeld, Marta  
APPLICANT: Ilya, Chumakov  
APPLICANT: Bouguetel, Lydie  
TITLE OF INVENTION: PROSTATE CANCER GENE  
FILE REFERENCE: GENSET.18CPICP  
CURRENT APPLICATION NUMBER: US/09/338,907  
EARLIER FILING DATE: 1999-06-23  
EARLIER APPLICATION NUMBER: 08/996,306  
EARLIER FILING DATE: 1997-12-22  
EARLIER APPLICATION NUMBER: 60/099,658  
EARLIER FILING DATE: 1998-09-09  
EARLIER APPLICATION NUMBER: 09/218,207  
EARLIER FILING DATE: 1998-12-22  
NUMBER OF SEQ ID NOS: 578  
SOFTWARE: Patent.pm  
SEQ ID NO 4  
LENGTH: 353  
TYPE: PRT  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: HELIX  
LOCATION: 1..33  
OTHER INFORMATION: Rao and Argos identification method, potential helix  
FEATURE:  
NAME/KEY: HELIX  
LOCATION: 4..20  
OTHER INFORMATION: Klein, Kanehisa and Delisi identification method, potent  
FEATURE:  
NAME/KEY: MYRISTATE  
LOCATION: 12..16  
OTHER INFORMATION: Prosite match  
FEATURE:  
NAME/KEY: HELIX  
LOCATION: 50..70  
OTHER INFORMATION: Eisenberg, Schwarz, Komarony, Wall identification method,  
FEATURE:  
NAME/KEY: CARBOHYD  
LOCATION: 57..59  
OTHER INFORMATION: Prosite match  
FEATURE:  
NAME/KEY: HELIX  
LOCATION: 76..96  
OTHER INFORMATION: Eisenberg, Schwarz, Komarony, Wall identification method,  
FEATURE:  
NAME/KEY: PHOSPHORYLATION  
LOCATION: 78  
OTHER INFORMATION: potential Tyrosine kinase site, Prosite match  
FEATURE:  
NAME/KEY: PHOSPHORYLATION  
LOCATION: 84  
OTHER INFORMATION: potential Tyrosine kinase II site, Prosite match  
FEATURE:  
NAME/KEY: SITE  
LOCATION: 94..115  
OTHER INFORMATION: potential Leucine zipper site, Prosite match  
FEATURE:  
NAME/KEY: MYRISTATE  
LOCATION: 119..123  
OTHER INFORMATION: potential site, Prosite match  
FEATURE:  
NAME/KEY: PHOSPHORYLATION  
LOCATION: 133  
OTHER INFORMATION: potential protein kinase C, Prosite match  
FEATURE:

```

: NAME/KEY: PHOSPHORYLATION
: LOCATION: 147
: OTHER INFORMATION: potential caseine kinase II site, Prosite match
: FEATURE:
: NAME/KEY: PHOSPHORYLATION
: LOCATION: 194
: OTHER INFORMATION: potential protein kinase C, Prosite match
: FEATURE:
: NAME/KEY: PHOSPHORYLATION
: LOCATION: 215
: OTHER INFORMATION: potential Tyrosine kinase site, Prosite match
: FEATURE:
: NAME/KEY: SULFATATION
: LOCATION: 221
: OTHER INFORMATION: Prosite match
: FEATURE:
: NAME/KEY: PHOSPHORYLATION
: LOCATION: 233
: OTHER INFORMATION: potential cAMP and cGMP dependant protein kinase site, Pr
: FEATURE:
: NAME/KEY: PHOSPHORYLATION
: LOCATION: 235
: OTHER INFORMATION: potential caseine kinase II site, Prosite match
: FEATURE:
: NAME/KEY: PHOSPHORYLATION
: LOCATION: 306
: OTHER INFORMATION: potential protein kinase C, Prosite match
: FEATURE:
: NAME/KEY: HELIX
: LOCATION: 310..330
: OTHER INFORMATION: Eisenberg, Schwarz, Komarony, Wall identification method, F
: FEATURE:
: NAME/KEY: MYRISTATE
: LOCATION: 319..323
: OTHER INFORMATION: Prosite match
: FEATURE:
: NAME/KEY: MYRISTATE
: LOCATION: 323..327
: OTHER INFORMATION: Prosite match
: FEATURE:
: NAME/KEY: AMIDATION
: LOCATION: 329
: OTHER INFORMATION: Prosite match
: FEATURE:
: NAME/KEY: HELIX
: LOCATION: 333..353
: OTHER INFORMATION: Eisenberg, Schwarz, Komarony, Wall identification method, F
: FEATURE:
: NAME/KEY: MYRISTATE
: LOCATION: 341..345
: OTHER INFORMATION: Prosite match
: FEATURE:
: NAME/KEY: PHOSPHORYLATION
: LOCATION: 350
: OTHER INFORMATION: potential protein kinase C, Prosite match
US-09-338-907-4
```

Query Match 100.0%; Score 228; DB 4; Length 353;

Best Local Similarity 100.0%; Pred. No. 1.6e-221; Indels 0; Gaps 0;

Matches 220; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

Qy 1 MRYLLPSVLLGTAFTYVLAHWRLSAPFLPARYQALDRLKCVQSMVLEFFENTY 60
Db 1 MRYLLPSVLLGTAFTYVLAHWRLSAPFLPARYQALDRLKCVQSMVLEFFENTY 60
Qy 61 VOILLYDLPKKNENIYYLANHSTVDWIVADILAIRONALGHVRYVLEGLKMLPLX 120
Db 61 VOILLYDLPKKNENIYYLANHSTVDWIVADILAIRONALGHVRYVLEGLKMLPLX 120
Qy 121 YFAOHGIIYVRSKAKFNKEMRKLQSVVDAGTSMYLVIFPEGGRYRNEQTKVLSASQAF 180
Db 121 YFAOHGIIYVRSKAKFNKEMRKLQSVVDAGTSMYLVIFPEGGRYRNEQTKVLSASQAF 180
```

```

Qy 181 AAGRLAVLKHVLPRIKATHVAFDCMKNYLDAIYDVTVYEGKDDG 228
Db 181 AAGRLAVLKHVLPRIKATHVAFDCMKNYLDAIYDVTVYEGKDDG 228
```

```

RESULT 5
US-09-218-207-4
: Sequence 4, Application US/09218207
: Patent No. 6346381
: GENERAL INFORMATION:
: APPLICANT: Cohen, Daniel
: APPLICANT: Blumenfeld, Maria
: APPLICANT: Ilya, Chumakov
: APPLICANT: Bougueleret, Lydie
: TITLE OF INVENTION: Prostate cancer gene
: FILE REFERENCE: GENSET 018CP1
: CURRENT APPLICATION NUMBER: US/09/218,207
: CURRENT FILING DATE: 1998-12-22
: EARLIER APPLICATION NUMBER: 08/996,306
: EARLIER FILING DATE: 1997-12-22
: EARLIER APPLICATION NUMBER: 60/099,658
: EARLIER FILING DATE: 1998-09-09
: NUMBER OF SEQ ID NOS: 578
: SOFTWARE: Patent.pm
: SEQ ID NO 4
: LENGTH: 353
: TYPE: PRT
: ORGANISM: Homo sapiens
: FEATURE:
: NAME/KEY: HELIX
: LOCATION: 1..33
: OTHER INFORMATION: Rao and Argos identification method, potential helix
: FEATURE:
: NAME/KEY: HELIX
: LOCATION: 4..20
: OTHER INFORMATION: Klein, Kanehisa and Delisi identification method, potent
: FEATURE:
: NAME/KEY: HELIX
: LOCATION: 4..24
: OTHER INFORMATION: Eisenberg, Schwarz, Komarony, Wall identification method,
: FEATURE:
: NAME/KEY: MYRISTATE
: LOCATION: 12..16
: OTHER INFORMATION: Prosite match
: FEATURE:
: NAME/KEY: HELIX
: LOCATION: 50..70
: OTHER INFORMATION: Eisenberg, Schwarz, Komarony, Wall identification method,
: FEATURE:
: NAME/KEY: CARBOHYD
: LOCATION: 57..59
: OTHER INFORMATION: Prosite match
: FEATURE:
: NAME/KEY: HELIX
: LOCATION: 76..96
: OTHER INFORMATION: Eisenberg, Schwarz, Komarony, Wall identification method,
: FEATURE:
: NAME/KEY: PHOSPHORYLATION
: LOCATION: 84
: OTHER INFORMATION: potential caseine kinase II site, Prosite match
: FEATURE:
: NAME/KEY: SITE
: LOCATION: 94..115
: OTHER INFORMATION: potential leucine zipper site, Prosite match
: FEATURE:
: NAME/KEY: MYRISTATE
: LOCATION: 119..123
: OTHER INFORMATION: potential site, Prosite match
```

FEATURE: PHOSPHORYLATION  
NAME/KEY: 133  
LOCATION: 133  
OTHER INFORMATION: potential protein kinase C, Prosite match  
FEATURE: PHOSPHORYLATION  
NAME/KEY: 147  
LOCATION: 147  
OTHER INFORMATION: potential caseine kinase II site, Prosite match  
FEATURE: PHOSPHORYLATION  
NAME/KEY: 194  
LOCATION: 194  
OTHER INFORMATION: potential protein kinase C, Prosite match  
FEATURE: PHOSPHORYLATION  
NAME/KEY: 215  
LOCATION: 215  
OTHER INFORMATION: potential Tyrosine kinase site, Prosite match  
FEATURE: SULFATATION  
NAME/KEY: 221  
LOCATION: 221  
OTHER INFORMATION: Prosite match  
FEATURE: PHOSPHORYLATION  
NAME/KEY: 223  
LOCATION: 223  
OTHER INFORMATION: potential cAMP and cGMP dependant protein kinase site, Prosite match  
FEATURE: PHOSPHORYLATION  
NAME/KEY: 235  
LOCATION: 235  
OTHER INFORMATION: potential caseine kinase II site, Prosite match  
FEATURE: PHOSPHORYLATION  
NAME/KEY: 306  
LOCATION: 306  
OTHER INFORMATION: potential protein kinase C, Prosite match  
FEATURE: HELIX  
NAME/KEY: 310..330  
LOCATION: 310..330  
OTHER INFORMATION: Eisenberg, Schwarz, Komarony, Wall identification method, F  
FEATURE: MYRISTATE  
NAME/KEY: 319..323  
LOCATION: 319..323  
OTHER INFORMATION: Prosite match  
FEATURE: MYRISTATE  
NAME/KEY: 323..327  
LOCATION: 323..327  
OTHER INFORMATION: Prosite match  
FEATURE: AMIDATION  
NAME/KEY: 329  
LOCATION: 329  
OTHER INFORMATION: Prosite match  
FEATURE: HELIX  
NAME/KEY: 333..353  
LOCATION: 333..353  
OTHER INFORMATION: Eisenberg, Schwarz, Komarony, Wall identification method, F  
FEATURE: MYRISTATE  
NAME/KEY: 341..345  
LOCATION: 341..345  
OTHER INFORMATION: Prosite match  
FEATURE: PHOSPHORYLATION  
NAME/KEY: 350  
LOCATION: 350  
OTHER INFORMATION: potential protein kinase C, Prosite match  
US-09-218-207-4

Query Match 100.0%; Score 228; DB 4; Length 353;  
Best Local Similarity 100.0%; Pred. No. 1.6e-221;  
Matches 228; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MRYLPSVLLGTAFTYVLAAGVWRLLSAFLPARFYQALDDRLCYVYQSWVLEFFENYTG 60  
|||||  
DB 1 MRYLPSVLLGTAFTYVLAAGVWRLLSAFLPARFYQALDDRLCYVYQSWVLEFFENYTG 60  
|||||  
QY 61 VOILLGDLPRKKNENIYLANHOSYDWMIVADILAIKRNALGHRVYVLEGLKWLPLCYG 120  
|||||

DB 61 VOILLGDLPRKKNENIYLANHOSYDWMIVADILAIKRNALGHRVYVLEGLKWLPLCYG 120  
QY 121 YFAOHGIIYKRSKAKNEKEMRNKLSYDAGTPTMYLVIPEGRVNPEDTKVLSASQAF 180  
|||||  
DB 121 YFAOHGIIYKRSKAKNEKEMRNKLSYDAGTPTMYLVIPEGRVNPEDTKVLSASQAF 180  
|||||  
QY 181 AARGIALVAKHVLTPRIKATHVAFDCMKNYLDIYDVTVVEGKDGG 228  
|||||  
DB 181 AARGIALVAKHVLTPRIKATHVAFDCMKNYLDIYDVTVVEGKDGG 228  
|||||

RESULT 6  
US-08-996-306-5  
; Sequence 5, Application US/08996306  
; Patent No. 5945522  
; GENERAL INFORMATION:  
; APPLICANT: Cohen, Daniel  
; APPLICANT: Chumakov, Ilya  
; APPLICANT: Blumenfeld, Marta  
; APPLICANT: Bougueniet, Lydie  
; TITLE OF INVENTION: Prostate cancer gene  
; NUMBER OF SEQUENCES: 68  
; CORRESPONDENCE ADDRESSES:  
; ADDRESSEE: Knobbe, Martens, Olson & Bear  
; STREET: 501 West Broadway  
; City: San Diego  
; STATE: California  
; COUNTRY: USA  
; ZIP: 92101-3505  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: Win95  
; SOFTWARE: word  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/996,306  
; FILING DATE:  
; CLASSIFICATION:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Israelien, Ned A.  
; REGISTRATION NUMBER: 29,655  
; REFERENCE/DOCKET NUMBER: GENSET.018A  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (619) 235-8550  
; TELEFAX: (619) 235-0176  
; INFORMATION FOR SEQ ID NO: 5:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 364 amino acids  
; TYPE: AMINO ACID  
; STRANDEDNESS: SINGLE  
; TOPOLOGY: LINEAR  
; MOLECULE TYPE: PROTEIN  
; ORIGINAL SOURCE:  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: potential protein  
; LOCATION: 1..364  
US-08-996-306-5

Query Match 100.0%; Score 228; DB 2; Length 364;  
Best Local Similarity 100.0%; Pred. No. 1.7e-221;  
Matches 228; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MRYLPSVLLGTAFTYVLAAGVWRLLSAFLPARFYQALDDRLCYVYQSWVLEFFENYTG 60  
|||||  
DB 12 MRYLPSVLLGTAFTYVLAAGVWRLLSAFLPARFYQALDDRLCYVYQSWVLEFFENYTG 71  
|||||  
QY 61 VOILLGDLPRKKNENIYLANHOSYDWMIVADILAIKRNALGHRVYVLEGLKWLPLCYG 120  
|||||  
DB 72 VOILLGDLPRKKNENIYLANHOSYDWMIVADILAIKRNALGHRVYVLEGLKWLPLCYG 131  
|||||  
QY 121 YFAOHGIIYKRSKAKNEKEMRNKLSYDAGTPTMYLVIPEGRVNPEDTKVLSASQAF 180  
|||||

```
Db 132 YFAOHGIIYKRSKAKFNEKEKRNKLSYVDAGTPMYLVIFPEGTRYNPEOTKVLASQAF 191
QY 181 AAORGLAVLKHVLPRIKATHVAFDCMKNYLDATYDVTYYEGKDDG 228
Db 192 AAORGLAVLKHVLPRIKATHVAFDCMKNYLDATYDVTYYEGKDDG 239
```

## RESULT 7

```
US-09-338-907-5
; Sequence 5, Application US/09338907
; Patent No. 6265546
; GENERAL INFORMATION:
; APPLICANT: Cohen, Daniel
; APPLICANT: Blumenfeld, Marta
; APPLICANT: Ilya, Chumakov
; APPLICANT: Bougueleret, Lydie
; TITLE OF INVENTION: PROSTATE CANCER GENE
; FILE REFERENCE: GENSET.18CPICP
; CURRENT APPLICATION NUMBER: US/09/338,907
; CURRENT FILING DATE: 1999-06-23
; EARLIER APPLICATION NUMBER: 08/996,306
; EARLIER FILING DATE: 1997-12-22
; EARLIER APPLICATION NUMBER: 60/099,658
; EARLIER FILING DATE: 1998-09-09
; EARLIER APPLICATION NUMBER: 09/218,207
; EARLIER FILING DATE: 1998-12-22
; NUMBER OF SEQ ID NOS: 578
; SOFTWARE: Patent.pm
; SEQ ID NO 5
; LENGTH: 364
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-338-907-5
```

```
Query Match 100.0%; Score 228; DB 4; Length 364;
Best Local Similarity 100.0%; Pred. No. 1.7e-221;
Matches 228; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 MRYLPSVVLGTPATPYVLAMGWRLSAPLPAFYQALDDRLCYQSVWLPFFENYTG 60
Db 12 MRYLPSVVLGTPATPYVLAMGWRLSAPLPAFYQALDDRLCYQSVWLPFFENYTG 71
QY 61 VOILLYDGLPRKNENIITYLANHOSTVDMIVADILAIQNALGHVRYVLEKGLKWLPLYGC 120
Db 72 VOILLYDGLPRKNENIITYLANHOSTVDMIVADILAIQNALGHVRYVLEKGLKWLPLYGC 131
QY 121 YFAOHGIIYKRSKAKFNEKEKRNKLSYVDAGTPMYLVIFPEGTRYNPEOTKVLASQAF 180
Db 132 YFAOHGIIYKRSKAKFNEKEKRNKLSYVDAGTPMYLVIFPEGTRYNPEOTKVLASQAF 191
QY 181 AAORGLAVLKHVLPRIKATHVAFDCMKNYLDATYDVTYYEGKDDG 228
Db 192 AAORGLAVLKHVLPRIKATHVAFDCMKNYLDATYDVTYYEGKDDG 239
```

## RESULT 8

```
US-09-218-207-5
; Sequence 5, Application US/09218207
; Patent No. 6346381
; GENERAL INFORMATION:
; APPLICANT: Cohen, Daniel
; APPLICANT: Blumenfeld, Marta
; APPLICANT: Ilya, Chumakov
; APPLICANT: Bougueleret, Lydie
; TITLE OF INVENTION: Prostate cancer gene
; FILE REFERENCE: GENSET.018CP1
; CURRENT APPLICATION NUMBER: US/09/218,207
; CURRENT FILING DATE: 1998-12-22
; EARLIER APPLICATION NUMBER: 08/996,306
; EARLIER FILING DATE: 1997-12-22
; EARLIER APPLICATION NUMBER: 60/099,658
```

```
; EARLIER FILING DATE: 1998-09-09
; NUMBER OF SEQ ID NOS: 578
; SOFTWARE: Patent.pm
; SEQ ID NO 5
; LENGTH: 364
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-218-207-5
```

```
Query Match 100.0%; Score 228; DB 4; Length 364;
Best Local Similarity 100.0%; Pred. No. 1.7e-221;
Matches 228; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 MRYLPSVVLGTPATPYVLAMGWRLSAPLPAFYQALDDRLCYQSVWLPFFENYTG 60
Db 12 MRYLPSVVLGTPATPYVLAMGWRLSAPLPAFYQALDDRLCYQSVWLPFFENYTG 71
QY 61 VOILLYDGLPRKNENIITYLANHOSTVDMIVADILAIQNALGHVRYVLEKGLKWLPLYGC 120
Db 72 VOILLYDGLPRKNENIITYLANHOSTVDMIVADILAIQNALGHVRYVLEKGLKWLPLYGC 131
QY 121 YFAOHGIIYKRSKAKFNEKEKRNKLSYVDAGTPMYLVIFPEGTRYNPEOTKVLASQAF 180
Db 132 YFAOHGIIYKRSKAKFNEKEKRNKLSYVDAGTPMYLVIFPEGTRYNPEOTKVLASQAF 191
QY 181 AAORGLAVLKHVLPRIKATHVAFDCMKNYLDATYDVTYYEGKDDG 228
Db 192 AAORGLAVLKHVLPRIKATHVAFDCMKNYLDATYDVTYYEGKDDG 239
```

## RESULT 9

```
US-09-338-907-136
; Sequence 136, Application US/09338907
; Patent No. 6265546
; GENERAL INFORMATION:
; APPLICANT: Cohen, Daniel
; APPLICANT: Blumenfeld, Marta
; APPLICANT: Ilya, Chumakov
; APPLICANT: Bougueleret, Lydie
; TITLE OF INVENTION: PROSTATE CANCER GENE
; FILE REFERENCE: GENSET.18CPICP
; CURRENT APPLICATION NUMBER: US/09/338,907
; CURRENT FILING DATE: 1999-06-23
; EARLIER APPLICATION NUMBER: 08/996,306
; EARLIER FILING DATE: 1997-12-22
; EARLIER APPLICATION NUMBER: 60/099,658
; EARLIER FILING DATE: 1998-09-09
; EARLIER APPLICATION NUMBER: 09/218,207
; EARLIER FILING DATE: 1998-12-22
; NUMBER OF SEQ ID NOS: 578
; SOFTWARE: Patent.pm
; SEQ ID NO 136
; LENGTH: 185
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: 81..83
; OTHER INFORMATION: Box I
; NAME/KEY: SITE
; LOCATION: 160..165
; OTHER INFORMATION: Box II
US-09-338-907-136
```

```
Query Match 81.1%; Score 185; DB 4; Length 185;
Best Local Similarity 100.0%; Pred. No. 1.6e-178;
Matches 185; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 MRYLPSVVLGTPATPYVLAMGWRLSAPLPAFYQALDDRLCYQSVWLPFFENYTG 60
|||||
```



```
Db      1 MRYLLPSVVLGTAPTYVLAMGWRLLSAFLPARFYQALDDRLCYCYQSMTLFEFFENYTG 60
QY      61 VOILLYGDLPRKKNENIYLANHOSYVDWIVADILAIRONALGHVRYVLEKGLKMLPLYGC 120
        |||
Db      61 VOILLYGDLPRKKNENIYLANHOSYVDWIVADILAIRONALGHVRYVLEKGLKMLPLYGC 120
QY      121 YFAOHGIGYKRSKAKFNEKEMRNKLOSVDAGTPLYVTFPEGSTRYNPOTKVLASQAF 180
        |||
Db      121 YFAOHGIGYKRSKAKFNEKEMRNKLOSVDAGTPLYVTFPEGSTRYNPOTKVLASQAF 180
QY      181 AAORG 185
        |||
Db      181 AAORG 185
```

```
RESULT 10
US-09-218-207-136
; Sequence 136, Application US/09218207
; Patent No. 6346381
; GENERAL INFORMATION:
; APPLICANT: Cohen, Daniel
; APPLICANT: Blumenfeld, Marta
; APPLICANT: Ilya, Chumakov
; APPLICANT: Bouqueleret, Lydie
; TITLE OF INVENTION: Prostate cancer gene
; FILE REFERENCE: GENSET.018CPI
; CURRENT APPLICATION NUMBER: US/09/218,207
; EARLIER FILING DATE: 1998-12-22
; EARLIER APPLICATION NUMBER: 08/996,306
; EARLIER FILING DATE: 1997-12-22
; EARLIER APPLICATION NUMBER: 60/099,658
; EARLIER FILING DATE: 1998-09-09
; NUMBER OF SEQ ID NOS: 578
; SOFTWARE: Patent.pm
; SEQ ID NO 136
; LENGTH: 185
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: 81..83
; OTHER INFORMATION: Box I
; FEATURE:
; NAME/KEY: SITE
; LOCATION: 160..165
; OTHER INFORMATION: Box II
US-09-218-207-136
```

```
Query Match      81.1%; Score 185; DB 4; Length 185;
Best Local Similarity 100.0%; Pred. No. 1.6e-178;
Matches 185; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      1 MRYLLPSVVLGTAPTYVLAMGWRLLSAFLPARFYQALDDRLCYCYQSMTLFEFFENYTG 60
        |||
Db      1 MRYLLPSVVLGTAPTYVLAMGWRLLSAFLPARFYQALDDRLCYCYQSMTLFEFFENYTG 60
QY      61 VOILLYGDLPRKKNENIYLANHOSYVDWIVADILAIRONALGHVRYVLEKGLKMLPLYGC 120
        |||
Db      61 VOILLYGDLPRKKNENIYLANHOSYVDWIVADILAIRONALGHVRYVLEKGLKMLPLYGC 120
QY      121 YFAOHGIGYKRSKAKFNEKEMRNKLOSVDAGTPLYVTFPEGSTRYNPOTKVLASQAF 180
        |||
Db      121 YFAOHGIGYKRSKAKFNEKEMRNKLOSVDAGTPLYVTFPEGSTRYNPOTKVLASQAF 180
QY      181 AAORG 185
        |||
Db      181 AAORG 185
```

```
RESULT 11
US-09-338-907-134
; Sequence 134, Application US/09338907
```

```
; Patent No. 6265546
; GENERAL INFORMATION:
; APPLICANT: Cohen, Daniel
; APPLICANT: Blumenfeld, Marta
; APPLICANT: Ilya, Chumakov
; APPLICANT: Bouqueleret, Lydie
; TITLE OF INVENTION: PROSTATE CANCER GENE
; FILE REFERENCE: GENSET.18CPICP
; CURRENT APPLICATION NUMBER: US/09/338,907
; EARLIER FILING DATE: 1999-06-23
; EARLIER APPLICATION NUMBER: 08/996,306
; EARLIER FILING DATE: 1997-12-22
; EARLIER APPLICATION NUMBER: 60/099,658
; EARLIER FILING DATE: 1998-09-09
; EARLIER APPLICATION NUMBER: 09/218,207
; EARLIER FILING DATE: 1998-12-22
; NUMBER OF SEQ ID NOS: 578
; SOFTWARE: Patent.pm
; SEQ ID NO 134
; LENGTH: 315
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: 81..83
; OTHER INFORMATION: Box I
; FEATURE:
; NAME/KEY: SITE
; LOCATION: 160..165
; OTHER INFORMATION: Box II
US-09-338-907-134
```

```
Query Match      81.1%; Score 185; DB 4; Length 315;
Best Local Similarity 100.0%; Pred. No. 2.7e-178;
Matches 185; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      1 MRYLLPSVVLGTAPTYVLAMGWRLLSAFLPARFYQALDDRLCYCYQSMTLFEFFENYTG 60
        |||
Db      1 MRYLLPSVVLGTAPTYVLAMGWRLLSAFLPARFYQALDDRLCYCYQSMTLFEFFENYTG 60
QY      61 VOILLYGDLPRKKNENIYLANHOSYVDWIVADILAIRONALGHVRYVLEKGLKMLPLYGC 120
        |||
Db      61 VOILLYGDLPRKKNENIYLANHOSYVDWIVADILAIRONALGHVRYVLEKGLKMLPLYGC 120
QY      121 YFAOHGIGYKRSKAKFNEKEMRNKLOSVDAGTPLYVTFPEGSTRYNPOTKVLASQAF 180
        |||
Db      121 YFAOHGIGYKRSKAKFNEKEMRNKLOSVDAGTPLYVTFPEGSTRYNPOTKVLASQAF 180
QY      181 AAORG 185
        |||
Db      181 AAORG 185
```

```
RESULT 12
US-09-218-207-134
; Sequence 134, Application US/09218207
; Patent No. 6346381
; GENERAL INFORMATION:
; APPLICANT: Cohen, Daniel
; APPLICANT: Blumenfeld, Marta
; APPLICANT: Ilya, Chumakov
; APPLICANT: Bouqueleret, Lydie
; TITLE OF INVENTION: Prostate cancer gene
; FILE REFERENCE: GENSET.018CPI
; CURRENT APPLICATION NUMBER: US/09/218,207
; EARLIER FILING DATE: 1998-12-22
; EARLIER APPLICATION NUMBER: 08/996,306
; EARLIER FILING DATE: 1997-12-22
; EARLIER APPLICATION NUMBER: 60/099,658
; EARLIER FILING DATE: 1998-09-09
; NUMBER OF SEQ ID NOS: 578
; SOFTWARE: Patent.pm
```

SEQ ID NO 134  
LENGTH: 315  
TYPE: prt  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: SITE  
LOCATION: 81..83  
OTHER INFORMATION: Box I  
FEATURE:  
NAME/KEY: SITE  
LOCATION: 160..165  
OTHER INFORMATION: Box II  
US-09-218-207-134

Query Match 81.1%; Score 185; DB 4; Length 315;  
Best Local Similarity 100.0%; Pred. No. 2.7e-178;  
Matches 185; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MRYLPSVVLGTAPTYVLAMGVWRLLSAFLPARFYQALDDRLCYQOSVWLEFFENYTG 60  
DB 1 MRYLPSVVLGTAPTYVLAMGVWRLLSAFLPARFYQALDDRLCYQOSVWLEFFENYTG 60  
QY 61 VOILLYGDLPKKNENITYLANHOSTVDWIVADILAIROMLGHVRYVLEKGLMWLPLYGC 120  
DB 61 VOILLYGDLPKKNENITYLANHOSTVDWIVADILAIROMLGHVRYVLEKGLMWLPLYGC 120  
QY 121 YFAOHGIIYKRSKAKNEKEMRNKLOSVDAGTPMVLVIFPEGTRYNPEOTKVLASQAF 180  
DB 121 YFAOHGIIYKRSKAKNEKEMRNKLOSVDAGTPMVLVIFPEGTRYNPEOTKVLASQAF 180  
QY 181 AAORG 185  
DB 181 AAORG 185

RESULT 13  
US-09-338-907-135  
Sequence 135, Application US/09338907  
Patent No. 6265546  
GENERAL INFORMATION:  
APPLICANT: Cohen, Daniel  
APPLICANT: Blumenfeld, Marla  
APPLICANT: Ilya, Chumakov  
APPLICANT: Bougueleret, Lydie  
TITLE OF INVENTION: PROSTATE CANCER GENE  
FILE REFERENCE: GENSET.18CPLP  
CURRENT APPLICATION NUMBER: US/09/338,907  
EARLIER FILING DATE: 1999-06-23  
EARLIER APPLICATION NUMBER: 08/996,306  
EARLIER FILING DATE: 1997-12-22  
EARLIER APPLICATION NUMBER: 60/099,658  
EARLIER FILING DATE: 1998-09-09  
EARLIER APPLICATION NUMBER: 09/218,207  
NUMBER OF SEQ ID NOS: 578  
SOFTWARE: Patent.pm  
SEQ ID NO 135  
LENGTH: 300  
TYPE: prt  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: SITE  
LOCATION: 81..83  
OTHER INFORMATION: Box I  
FEATURE:  
NAME/KEY: SITE  
LOCATION: 160..165  
OTHER INFORMATION: Box II  
US-09-338-907-135

Query Match 80.7%; Score 184; DB 4; Length 300;

Best Local Similarity 100.0%; Pred. No. 2.6e-177;  
Matches 184; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MRYLPSVVLGTAPTYVLAMGVWRLLSAFLPARFYQALDDRLCYQOSVWLEFFENYTG 60  
DB 1 MRYLPSVVLGTAPTYVLAMGVWRLLSAFLPARFYQALDDRLCYQOSVWLEFFENYTG 60  
QY 61 VOILLYGDLPKKNENITYLANHOSTVDWIVADILAIROMLGHVRYVLEKGLMWLPLYGC 120  
DB 61 VOILLYGDLPKKNENITYLANHOSTVDWIVADILAIROMLGHVRYVLEKGLMWLPLYGC 120  
QY 121 YFAOHGIIYKRSKAKNEKEMRNKLOSVDAGTPMVLVIFPEGTRYNPEOTKVLASQAF 180  
DB 121 YFAOHGIIYKRSKAKNEKEMRNKLOSVDAGTPMVLVIFPEGTRYNPEOTKVLASQAF 180  
QY 181 AAOR 184  
DB 181 AAOR 184

RESULT 14  
US-09-218-207-135  
Sequence 135, Application US/09218207  
Patent No. 6346381  
GENERAL INFORMATION:  
APPLICANT: Cohen, Daniel  
APPLICANT: Blumenfeld, Marla  
APPLICANT: Ilya, Chumakov  
APPLICANT: Bougueleret, Lydie  
TITLE OF INVENTION: Prostate cancer gene  
FILE REFERENCE: GENSET.018CPL  
CURRENT APPLICATION NUMBER: US/09/218,207  
EARLIER FILING DATE: 1998-12-22  
EARLIER APPLICATION NUMBER: 08/996,306  
EARLIER FILING DATE: 1997-12-22  
EARLIER APPLICATION NUMBER: 60/099,658  
EARLIER FILING DATE: 1998-09-09  
NUMBER OF SEQ ID NOS: 578  
SOFTWARE: Patent.pm  
SEQ ID NO 135  
LENGTH: 300  
TYPE: prt  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: SITE  
LOCATION: 81..83  
OTHER INFORMATION: Box I  
FEATURE:  
NAME/KEY: SITE  
LOCATION: 160..165  
OTHER INFORMATION: Box II  
US-09-218-207-135

Query Match 80.7%; Score 184; DB 4; Length 300;  
Best Local Similarity 100.0%; Pred. No. 2.6e-177;  
Matches 184; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MRYLPSVVLGTAPTYVLAMGVWRLLSAFLPARFYQALDDRLCYQOSVWLEFFENYTG 60  
DB 1 MRYLPSVVLGTAPTYVLAMGVWRLLSAFLPARFYQALDDRLCYQOSVWLEFFENYTG 60  
QY 61 VOILLYGDLPKKNENITYLANHOSTVDWIVADILAIROMLGHVRYVLEKGLMWLPLYGC 120  
DB 61 VOILLYGDLPKKNENITYLANHOSTVDWIVADILAIROMLGHVRYVLEKGLMWLPLYGC 120  
QY 121 YFAOHGIIYKRSKAKNEKEMRNKLOSVDAGTPMVLVIFPEGTRYNPEOTKVLASQAF 180  
DB 121 YFAOHGIIYKRSKAKNEKEMRNKLOSVDAGTPMVLVIFPEGTRYNPEOTKVLASQAF 180  
QY 181 AAOR 184  
DB 181 AAOR 184

RESULT 15  
US-09-338-907-133  
; Sequence 133, Application US/09338907  
; Patent No. 6265546  
; GENERAL INFORMATION:  
; APPLICANT: Cohen, Daniel  
; APPLICANT: Blumenfeld, Marta  
; APPLICANT: Ilya, Chumakov  
; APPLICANT: Bouquelerec, Lydie  
; TITLE OF INVENTION: PROSTATE CANCER GENE  
; FILE REFERENCE: GENSET.18C1CP  
; CURRENT APPLICATION NUMBER: us/09/338, 907  
; CURRENT FILING DATE: 1999-06-23  
; EARLIER APPLICATION NUMBER: 08/996, 306  
; EARLIER FILING DATE: 1997-12-22  
; EARLIER APPLICATION NUMBER: 60/099, 658  
; EARLIER FILING DATE: 1998-09-09  
; EARLIER APPLICATION NUMBER: 09/218, 207  
; EARLIER FILING DATE: 1998-12-22  
; NUMBER OF SEQ ID NOS: 578  
; SOFTWARE: Patent.pm  
; SEQ ID NO 133  
; LENGTH: 182  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: SITE  
; LOCATION: 81..83  
; OTHER INFORMATION: Box I  
US-09-338-907-133

Query Match 67.58; Score 154; DB 4; Length 182;  
Best Local Similarity 100.0%; Pred. No. 2.5e-147;  
Matches 154; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 MRYLLPSVVLGTAPTYVLAWGVWRLLSAFLPARFYQALDDRLCYQOSMVLFFENYTG 60  
Db 1 MRYLLPSVVLGTAPTYVLAWGVWRLLSAFLPARFYQALDDRLCYQOSMVLFFENYTG 60  
QY 61 VOILLYGDLPKKNENIYLANHOSVDWIVADILAIRONALGHVRYVLKEGLKWLPLYGC 120  
Db 61 VOILLYGDLPKKNENIYLANHOSVDWIVADILAIRONALGHVRYVLKEGLKWLPLYGC 120  
QY 121 YFAOHGIIYVKRSKAFNEKEMRNKLQSYVDAGTP 154  
Db 121 YFAOHGIIYVKRSKAFNEKEMRNKLQSYVDAGTP 154

Search completed: August 28, 2002, 11:19:22  
Job time: 305 sec

**THIS PAGE BLANK (USPTO)**

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 28, 2002, 11:25:19 ; Search time 305.3 Seconds

(without alignments)  
262.861 Million cell updates/sec

Title: US-09-853-526-70

Perfect score: 228

Sequence: 1 MRYLPSVVLGTAPTYVLA.....NYLDATYVTVEYKDDGG 228

Scoring table:

Gapop 60.0, Gapext 60.0

Searched: 3502263 seqs, 351980561 residues

Word size : 0

Total number of hits satisfying chosen parameters: 3502263

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : Pending\_Patents\_AA\_Main:

1: /cgn2\_6/ptodata/2/paa/PCYRUS.COMB.pep:\*  
2: /cgn2\_6/ptodata/2/paa/US06.COMB.pep:\*  
3: /cgn2\_6/ptodata/2/paa/US07.COMB.pep:\*  
4: /cgn2\_6/ptodata/2/paa/US08.COMB.pep:\*  
5: /cgn2\_6/ptodata/2/paa/US081.COMB.pep:\*  
6: /cgn2\_6/ptodata/2/paa/US082.COMB.pep:\*  
7: /cgn2\_6/ptodata/2/paa/US083.COMB.pep:\*  
8: /cgn2\_6/ptodata/2/paa/US084.COMB.pep:\*  
9: /cgn2\_6/ptodata/2/paa/US085.COMB.pep:\*  
10: /cgn2\_6/ptodata/2/paa/US086.COMB.pep:\*  
11: /cgn2\_6/ptodata/2/paa/US087.COMB.pep:\*  
12: /cgn2\_6/ptodata/2/paa/US088.COMB.pep:\*  
13: /cgn2\_6/ptodata/2/paa/US089.COMB.pep:\*  
14: /cgn2\_6/ptodata/2/paa/US090.COMB.pep:\*  
15: /cgn2\_6/ptodata/2/paa/US091.COMB.pep:\*  
16: /cgn2\_6/ptodata/2/paa/US092.COMB.pep:\*  
17: /cgn2\_6/ptodata/2/paa/US093.COMB.pep:\*  
18: /cgn2\_6/ptodata/2/paa/US094.COMB.pep:\*  
19: /cgn2\_6/ptodata/2/paa/US095.COMB.pep:\*  
20: /cgn2\_6/ptodata/2/paa/US096.COMB.pep:\*  
21: /cgn2\_6/ptodata/2/paa/US097.COMB.pep:\*  
22: /cgn2\_6/ptodata/2/paa/US098.COMB.pep:\*  
23: /cgn2\_6/ptodata/2/paa/US099.COMB.pep:\*  
24: /cgn2\_6/ptodata/2/paa/US100.COMB.pep:\*  
25: /cgn2\_6/ptodata/2/paa/US101.COMB.pep:\*  
26: /cgn2\_6/ptodata/2/paa/US60.COMB.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	228	100.0	228	22	US-09-853-526-70
2	228	100.0	228	23	US-09-901-484A-70
3	228	100.0	228	26	US-60-099-658-70
4	228	100.0	353	1	PCT-US01-01431-59
5	228	100.0	353	1	PCT-US01-01431-79
6	228	100.0	353	1	PCT-US01-11988-1474
7	228	100.0	353	1	PCT-US01-11988-1475

8	228	100.0	353	18	US-09-488-725A-2736	Sequence 2736, Ap
9	228	100.0	353	22	US-09-833-245-1474	Sequence 1474, Ap
10	228	100.0	353	22	US-09-833-245-1475	Sequence 1475, Ap
11	228	100.0	353	22	US-09-853-526-4	Sequence 4, Appl1
12	228	100.0	353	23	US-09-901-464A-4	Sequence 4, Appl1
13	228	100.0	353	23	US-09-915-582-59	Sequence 59, Appl1
14	228	100.0	353	23	US-09-915-582-79	Sequence 79, Appl1
15	228	100.0	353	26	US-60-099-658-4	Sequence 4, Appl1
16	228	100.0	364	18	US-09-436-919-1	Sequence 1, Appl1
17	228	100.0	364	22	US-09-817-910-7	Sequence 7, Appl1
18	228	100.0	364	22	US-09-853-526-5	Sequence 5, Appl1
19	228	100.0	364	23	US-09-901-484A-5	Sequence 5, Appl1
20	228	100.0	364	26	US-60-099-658-5	Sequence 5, Appl1
21	185	81.1	185	22	US-09-853-526-136	Sequence 136, App
22	185	81.1	185	23	US-09-901-484A-136	Sequence 136, App
23	185	81.1	315	22	US-09-853-526-134	Sequence 134, App
24	185	81.1	315	23	US-09-901-484A-134	Sequence 134, App
25	184	80.7	300	22	US-09-853-526-135	Sequence 135, App
26	184	80.7	300	23	US-09-901-484A-135	Sequence 135, App
27	154	67.5	182	22	US-09-853-526-133	Sequence 133, App
28	154	67.5	182	23	US-09-901-484A-133	Sequence 133, App
29	144	63.2	269	1	PCT-US01-01327-47	Sequence 47, Appl
30	143	62.7	450	26	US-60-245-222-110	Sequence 110, App
31	142	62.3	1032	26	US-60-212-413-309	Sequence 309, App
32	142	62.3	1032	26	US-60-229-518-248	Sequence 248, App
33	127	55.7	372	18	US-09-488-725A-6308	Sequence 6308, App
34	105	46.1	291	22	US-09-853-526-127	Sequence 127, App
35	105	46.1	291	23	US-09-901-484A-127	Sequence 127, App
36	91	39.9	176	18	US-09-436-919-5	Sequence 5, Appl1
37	85	37.3	97	22	US-09-853-526-132	Sequence 132, App
38	85	37.3	97	23	US-09-901-484A-132	Sequence 132, App
39	81	35.5	92	17	US-09-316-123-124	Sequence 124, App
40	81	35.5	92	22	US-09-827-244-124	Sequence 124, App
41	74	32.5	261	22	US-09-853-526-128	Sequence 128, App
42	74	32.5	261	23	US-09-901-484A-128	Sequence 128, App
43	70	30.7	257	1	PCT-US01-08631-31719	Sequence 31719, A
44	70	30.7	980	1	PCT-US01-08631-42336	Sequence 42336, A
45	62	27.2	66	22	US-09-853-526-131	Sequence 131, App

#### ALIGNMENTS

RESULT 1  
US-09-853-526-70  
; Sequence 70, Application US/09853526  
; GENERAL INFORMATION:  
; APPLICANT: Cohen, Daniel  
; APPLICANT: Blumenfeld, Marta  
; APPLICANT: Ilya, Chumakov  
; APPLICANT: Bougueleret, Lydie  
; TITLE OF INVENTION: PROSTATE CANCER GENE  
; FILE REFERENCE: GENSET.18CPICP  
; CURRENT APPLICATION NUMBER: US/09/853,526  
; CURRENT FILING DATE: 2001-05-11  
; PRIOR APPLICATION NUMBER: 09/338,907  
; PRIOR FILING DATE: 1999-06-23  
; PRIOR APPLICATION NUMBER: 08/996,306  
; PRIOR FILING DATE: 1997-12-22  
; PRIOR APPLICATION NUMBER: 60/099,658  
; PRIOR FILING DATE: 1998-09-09  
; PRIOR APPLICATION NUMBER: 09/218,207  
; PRIOR FILING DATE: 1998-12-22  
; NUMBER OF SEQ ID NOS: 578  
; SOFTWARE: Patent.pm  
; SEQ ID NO 70  
; LENGTH: 228  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-853-526-70  
Query Match 100.0%; Score 228; DB 22; Length 228;

Best Local Similarity 100.0%; Pred. NO. 6.3e-221;  
Matches 228; Conservative 0; Mismatches 0; Indels 0;  
Gaps 0

Qy	1	MRLLBSVLLGTAAPYVLANGVWLLSAFLPAFPAFYALDRLCYQASVWLFEEFNNTYG	60
Db	1	MRLLBSVLLGTAAPYVLANGVWLLSAFLPAFPAFYALDRLCYQASVWLFEEFNNTYG	60
Qy	61	VQILLGVDDLPRKNENITVLANHSTVDVIMVADILAIROMALGHRVYLLEGKLMPLCYGC	12
Db	61	VQILLGVDDLPRKNENITVLANHSTVDVIMVADILAIROMALGHRVYLLEGKLMPLCYGC	12
Qy	121	YFQDHGGIYVKKSAFNEKEARNKIKQSTVVDAGTMYLYIFPEEGTRINPEQTKYLSAQAF	18
Db	121	YFQDHGGIYVKKSAFNEKEARNKIKQSTVVDAGTMYLYIFPEEGTRINPEQTKYLSAQAF	18
Qy	181	AAQRGLAVLKHVLTPIRIKTHVAFQCKMKNYDAILYDVTYVYEGGDGG	228
Db	181	AAQRGLAVLKHVLTPIRIKTHVAFQCKMKNYDAILYDVTYVYEGGDGG	228

## RESULT 2

```

US-09-901-484A-70
: Sequence 70. Application US/09901484A
: GENERAL INFORMATION:
: APPLICANT: Blumen, Daniel
: APPLICANT: Blumenfeld, Marta
: APPLICANT: Chumakov, Ilya
: APPLICANT: Bouguenelat, Lydie
: TITLE OF INVENTION: Prostate Cancer Gene
: FILE REFERENCE: GEN-1111XC302
: CURRENT APPLICATION NUMBER: US/09/901,484A
: PRIORITY FILING DATE: 2001-07-09
: PRIOR APPLICATION NUMBER: US 08/996,306
: PRIOR FILING DATE: 1997-12-22
: PRIOR APPLICATION NUMBER: US 60/099,658
: PRIOR FILING DATE: 1998-09-09
: PRIOR APPLICATION NUMBER: US 09/218,207
: PRIOR FILING DATE: 1998-12-22
: PRIOR APPLICATION NUMBER: US 09/338,907
: PRIOR FILING DATE: 1999-06-23
: PRIOR APPLICATION NUMBER: US 09/853,526
: PRIOR FILING DATE: 2001-05-11
: NUMBER OF SEQ ID NOS: 578
: SOFTWARE: PatentIn version 3.1
: SEQ ID NO 70
: LENGTH: 228
: TYPE: PRT
: ORGANISM: Homo sapiens
: FEATURE:
: NAME/KEY: misc_feature
: LOCATION: (5208)..(5208)
: OTHER INFORMATION: n = a, c, g, or t.
US-09-901-484A-70

```

Query Match	100.0%	Score 228	DB 23	Length 228
-------------	--------	-----------	-------	------------

	Matches	228:	Conservative	0:	Mismatches	0:	Indels	0:	Gaps	0:
OY	1	MRLLPSVVLGTAFTYYLAWGWRLLSAFLEPARYQALDRLCYQSOMLVFFENNYTG	60							
Db	1	MRLLPSVVLGTAFTYYLAWGWRLLSAFLEPARYQALDRLCYQSOMLVFFENNYTG	60							
OY	61	VQILLGLDPKNNKEIITLANHGSTVDIMIVADILAIROMALGHVRVYLKEGIKLWPLGYC	120							
Db	61	VQILLGLDPKNNKEIITLANHGSTVDIMIVADILAIROMALGHVRVYLKEGIKLWPLGYC	120							
OY	121	YFAOHGGIYVRSAKFENKEMRNKIQSYVDAGTPMYLVTFEPGSTRYNPEQTLYLASQAQF	180							
Db	121	YFAOHGGIYVRSAKFENKEMRNKIQSYVDAGTPMYLVTFEPGSTRYNPEQTLYLASQAQF	180							
OY	181	AAGGGLAVLKHLVLPRIKATHYAIDCMKNYLDATIDVYVYVEGSKDDG	228							
Db	181	AAGGGLAVLKHLVLPRIKATHYAIDCMKNYLDATIDVYVYVEGSKDDG	228							

Db 181 AAQGLAVLKHVLTPIKATHVAFDCMKNYLDAIYDVTVYEGKDDG 228

RESULT 3

```

Sequence 70, Application US/60093658
GENERAL INFORMATION:
APPLICANT: Cohen, Daniel
APPLICANT: Chumakov, Ilya
APPLICANT: Blumenfeld, Marta
APPLICANT: Bouguenieret, Lydie
TITLE OF INVENTION: Prostate cancer gene
NUMBER OF SEQUENCES: 99
CORRESPONDENCE ADDRESS:
ADDRESSEE: Knobbe, Martens, Olson & Bear
STREET: 501 West Broadway
CITY: San Diego
STATE: California
COUNTRY: USA
ZIP: 92101-3505
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy Disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: Mln95
SOFTWARE: Word
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/60/099,658
FILING DATE:
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Israelson, Ned A.
REGISTRATION NUMBER: 29,655
REFERENCE/DOCKET NUMBER: GENSET.018APR
TELECOMMUNICATION INFORMATION:
TELEPHONE: (619) 235-8550
TELEFAX: (619) 235-0176
INFORMATION FOR SEQ ID NO: 70:
SEQUENCE CHARACTERISTICS:
LENGTH: 228 amino acids
TYPE: AMINO ACID
STRANDEDNESS: SINGLE
TOPOLOGY: LINEAR
MOLECULE TYPE: Protein
ORIGINAL SOURCE:
ORGANISM: Homo sapiens
IS-60-099-658-70

```

query match	100.0%	Score 228;	DB 26;	Length 228;
-------------	--------	------------	--------	-------------

	Matches	228	Conservative	0	Mismatches	0	Indels	0	Gaps
Qy	1	MRYLTPSVLLIGTAPTYVLAMGVWRLLSAPLPAFYQALDDRLCYQVQSNVLEFFENYTG	60						
Db	1	MRYLTPSVLLIGTAPTYVLAMGVWRLLSAPLPAFYQALDDRLCYQVQSNVLEFFENYTG	60						
Qy	61	VQILLYDDELPRNKENIITYLANHOSYVDIMYADLIAIQONALGHRVYLKSGKMLPLYGC	120						
Db	61	VQILLYDDELPRNKENIITYLANHOSYVDIMYADLIAIQONALGHRVYLKSGKMLPLYGC	120						
Qy	121	YFAOHGCIYKRSKAFNEKEKEMRNKLOSYVDAGTPMYLVIFPEGRIRNPEDTKVLSAQAF	180						
Db	121	YFAOHGCIYKRSKAFNEKEKEMRNKLOSYVDAGTPMYLVIFPEGRIRNPEDTKVLSAQAF	180						
Qy	181	AAOAGLAVLKHVLTPIKATHVAAPDCCKNVLADYDTVTYVEGGDDSG	228						
Db	181	AAOAGLAVLKHVLTPIKATHVAAPDCCKNVLADYDTVTYVEGGDDSG	228						

## RESULT 4

```

PCT-US01-01431-59
; Sequence 59, Application PC/TUS0101431
; GENERAL INFORMATION:

```

```
; APPLICANT: Human Genome Sciences, Inc.
; TITLE OF INVENTION: 17 human secreted proteins
; FILE REFERENCE: PS723PCT
; CURRENT APPLICATION NUMBER: PCT/US01/01431
; CURRENT FILING DATE: 2001-01-17
; PRIOR APPLICATION NUMBER: 60/179,065
; PRIOR FILING DATE: 2000-01-31
; PRIOR APPLICATION NUMBER: 60/180,628
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: 60/231,968
; PRIOR FILING DATE: 2000-09-12
; NUMBER OF SEQ ID NOS: 97
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 59
; LENGTH: 353
; TYPE: PRT
; ORGANISM: Homo sapiens
PCT-US01-01431-59
```

```
Query Match          100.0%; Score 228; DB 1; Length 353;
Best Local Similarity 100.0%; Pred. No. 9.3e-221;
Matches 228; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 MRYLLPSVVLGTAPTYVLAMGWRLLSAFLPARFYQALDDRLCYQYSQWVLEFFENYTG 60
DB 1 MRYLLPSVVLGTAPTYVLAMGWRLLSAFLPARFYQALDDRLCYQYSQWVLEFFENYTG 60
QY 61 VOILLYGDLPKKNENIYYLANHOSYDWMIVADILAIRONALGHVRYVLEKGLKWLPLYGC 120
DB 61 VOILLYGDLPKKNENIYYLANHOSYDWMIVADILAIRONALGHVRYVLEKGLKWLPLYGC 120
QY 121 YFAOHGGIYVKSARENEKEMRNKLOSVDAGTPTLVIFPESTRNPEQTKVLSASQAF 180
DB 121 YFAOHGGIYVKSARENEKEMRNKLOSVDAGTPTLVIFPESTRNPEQTKVLSASQAF 180
QY 181 AAORGLAVLKHVLTPIKATHVAFDCKMKNYLDIYDVTVVYEESKDDGG 228
DB 181 AAORGLAVLKHVLTPIKATHVAFDCKMKNYLDIYDVTVVYEESKDDGG 228
```

```
RESULT 5
PCT-US01-01431-79
; Sequence 79, Application PC/TUS0101431
; GENERAL INFORMATION:
; APPLICANT: Human Genome Sciences, Inc.
; TITLE OF INVENTION: 17 human secreted proteins
; FILE REFERENCE: PS723PCT
; CURRENT APPLICATION NUMBER: PCT/US01/01431
; CURRENT FILING DATE: 2001-01-17
; PRIOR APPLICATION NUMBER: 60/179,065
; PRIOR FILING DATE: 2000-01-31
; PRIOR APPLICATION NUMBER: 60/180,628
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: 60/231,968
; PRIOR FILING DATE: 2000-09-12
; NUMBER OF SEQ ID NOS: 97
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 79
; LENGTH: 353
; TYPE: PRT
; ORGANISM: Homo sapiens
PCT-US01-01431-79
```

```
Query Match          100.0%; Score 228; DB 1; Length 353;
Best Local Similarity 100.0%; Pred. No. 9.3e-221;
Matches 228; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 MRYLLPSVVLGTAPTYVLAMGWRLLSAFLPARFYQALDDRLCYQYSQWVLEFFENYTG 60
DB 1 MRYLLPSVVLGTAPTYVLAMGWRLLSAFLPARFYQALDDRLCYQYSQWVLEFFENYTG 60
```

```
QY 61 VOILLYGDLPKKNENIYYLANHOSYDWMIVADILAIRONALGHVRYVLEKGLKWLPLYGC 120
DB 61 VOILLYGDLPKKNENIYYLANHOSYDWMIVADILAIRONALGHVRYVLEKGLKWLPLYGC 120
QY 121 YFAOHGGIYVKSARENEKEMRNKLOSVDAGTPTLVIFPESTRNPEQTKVLSASQAF 180
DB 121 YFAOHGGIYVKSARENEKEMRNKLOSVDAGTPTLVIFPESTRNPEQTKVLSASQAF 180
QY 181 AAORGLAVLKHVLTPIKATHVAFDCKMKNYLDIYDVTVVYEESKDDGG 228
DB 181 AAORGLAVLKHVLTPIKATHVAFDCKMKNYLDIYDVTVVYEESKDDGG 228
```

```
RESULT 6
PCT-US01-11988-1474
; Sequence 1474, Application PC/TUS0111988
; GENERAL INFORMATION:
; APPLICANT: Human Genome Sciences, Inc.
; TITLE OF INVENTION: Albumin Fusion Proteins
; FILE REFERENCE: PF546PCT
; CURRENT APPLICATION NUMBER: PCT/US01/11988
; CURRENT FILING DATE: 2001-01-12
; PRIOR APPLICATION NUMBER: 60/229,358
; PRIOR FILING DATE: 2000-04-12
; PRIOR APPLICATION NUMBER: 60/256,931
; PRIOR FILING DATE: 2000-12-21
; PRIOR APPLICATION NUMBER: 60/199,384
; PRIOR FILING DATE: 2000-04-25
; NUMBER OF SEQ ID NOS: 2267
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 1474
; LENGTH: 353
; TYPE: PRT
; ORGANISM: Homo sapiens
PCT-US01-11988-1474
```

```
Query Match          100.0%; Score 228; DB 1; Length 353;
Best Local Similarity 100.0%; Pred. No. 9.3e-221;
Matches 228; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 MRYLLPSVVLGTAPTYVLAMGWRLLSAFLPARFYQALDDRLCYQYSQWVLEFFENYTG 60
DB 1 MRYLLPSVVLGTAPTYVLAMGWRLLSAFLPARFYQALDDRLCYQYSQWVLEFFENYTG 60
QY 61 VOILLYGDLPKKNENIYYLANHOSYDWMIVADILAIRONALGHVRYVLEKGLKWLPLYGC 120
DB 61 VOILLYGDLPKKNENIYYLANHOSYDWMIVADILAIRONALGHVRYVLEKGLKWLPLYGC 120
QY 121 YFAOHGGIYVKSARENEKEMRNKLOSVDAGTPTLVIFPESTRNPEQTKVLSASQAF 180
DB 121 YFAOHGGIYVKSARENEKEMRNKLOSVDAGTPTLVIFPESTRNPEQTKVLSASQAF 180
QY 181 AAORGLAVLKHVLTPIKATHVAFDCKMKNYLDIYDVTVVYEESKDDGG 228
DB 181 AAORGLAVLKHVLTPIKATHVAFDCKMKNYLDIYDVTVVYEESKDDGG 228
```

```
RESULT 7
PCT-US01-11988-1475
; Sequence 1475, Application PC/TUS0111988
; GENERAL INFORMATION:
; APPLICANT: Human Genome Sciences, Inc.
; TITLE OF INVENTION: Albumin Fusion Proteins
; FILE REFERENCE: PF546PCT
; CURRENT APPLICATION NUMBER: PCT/US01/11988
; CURRENT FILING DATE: 2001-01-12
; PRIOR APPLICATION NUMBER: 60/229,358
; PRIOR FILING DATE: 2000-04-12
; PRIOR APPLICATION NUMBER: 60/256,931
; PRIOR FILING DATE: 2000-12-21
; PRIOR APPLICATION NUMBER: 60/199,384
; PRIOR FILING DATE: 2000-04-25
```

NUMBER OF SEQ ID NOS: 2267  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 1475  
LENGTH: 353  
TYPE: PRT  
ORGANISM: Homo sapiens  
PCT-US01-11988-1475

Query Match 100.0%; Score 228; DB 1; Length 353;  
Best Local Similarity 100.0%; Pred. No. 9.3e-221;  
Matches 228; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MRYLLPSVLLGTAPTYVLAAGVWRLLSAFLEPAFYQALDRLCYQSVWLEFFENYTG 60  
DB 1 MRYLLPSVLLGTAPTYVLAAGVWRLLSAFLEPAFYQALDRLCYQSVWLEFFENYTG 60  
QY 61 VOILLYDGLPKNKENIITYLANHSTVDMIVADILAIQNALGHVRYVLEKGLKPLPGC 120  
DB 61 VOILLYDGLPKNKENIITYLANHSTVDMIVADILAIQNALGHVRYVLEKGLKPLPGC 120  
QY 121 YFAOHGIIYVRSKAKFENEKERNKLOSYVDAGTPMYLVIFPEGTRYNPEOTKVLASQAF 180  
DB 121 YFAOHGIIYVRSKAKFENEKERNKLOSYVDAGTPMYLVIFPEGTRYNPEOTKVLASQAF 180  
QY 181 AAOGLAVLKHVLTLPRIKATHVAFDCMKNYLDALYDVTVYEGKDDG 228  
DB 181 AAOGLAVLKHVLTLPRIKATHVAFDCMKNYLDALYDVTVYEGKDDG 228

RESULT 8  
US-09-488-725A-2736  
Sequence 2736, Application US/09488725A  
GENERAL INFORMATION:  
APPLICANT: Hyseq Inc  
TITLE OF INVENTION: Novel Nucleic Acid and Polypeptides  
FILE REFERENCE: 784FDPCT  
CURRENT APPLICATION NUMBER: US/09/488,725A  
CURRENT FILING DATE: 2000-12-22  
PRIOR APPLICATION NUMBER: US/09/488,725  
PRIOR FILING DATE: 2000-01-21  
PRIOR APPLICATION NUMBER: US09/552,317  
PRIOR FILING DATE: 2000-04-25  
PRIOR APPLICATION NUMBER: US09/598,042  
PRIOR FILING DATE: 2000-06-20  
PRIOR APPLICATION NUMBER: US09/620,312  
PRIOR FILING DATE: 2000-07-19  
PRIOR APPLICATION NUMBER: US09/653,450  
PRIOR FILING DATE: 2000-08-31  
PRIOR APPLICATION NUMBER: US09/662,191  
PRIOR FILING DATE: 2000-09-14  
PRIOR APPLICATION NUMBER: US09/693,036  
PRIOR FILING DATE: 2000-10-19  
PRIOR APPLICATION NUMBER: US09/727,344  
PRIOR FILING DATE: 2000-11-29  
NUMBER OF SEQ ID NOS: 7144  
SOFTWARE: PL\_FL\_genes\_b Versions 1.0  
SEQ ID NO 2736  
LENGTH: 354  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-09-488-725A-2736

Query Match 100.0%; Score 228; DB 18; Length 353;  
Best Local Similarity 100.0%; Pred. No. 9.3e-221;  
Matches 228; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MRYLLPSVLLGTAPTYVLAAGVWRLLSAFLEPAFYQALDRLCYQSVWLEFFENYTG 60  
DB 1 MRYLLPSVLLGTAPTYVLAAGVWRLLSAFLEPAFYQALDRLCYQSVWLEFFENYTG 60  
QY 61 VOILLYDGLPKNKENIITYLANHSTVDMIVADILAIQNALGHVRYVLEKGLKPLPGC 120

DB 61 VOILLYDGLPKNKENIITYLANHSTVDMIVADILAIQNALGHVRYVLEKGLKPLPGC 120  
QY 121 YFAOHGIIYVRSKAKFENEKERNKLOSYVDAGTPMYLVIFPEGTRYNPEOTKVLASQAF 180  
DB 121 YFAOHGIIYVRSKAKFENEKERNKLOSYVDAGTPMYLVIFPEGTRYNPEOTKVLASQAF 180  
QY 181 AAOGLAVLKHVLTLPRIKATHVAFDCMKNYLDALYDVTVYEGKDDG 228  
DB 181 AAOGLAVLKHVLTLPRIKATHVAFDCMKNYLDALYDVTVYEGKDDG 228

RESULT 9  
US-09-833-245-1474  
Sequence 1474, Application US/09833245  
GENERAL INFORMATION:  
APPLICANT: Human Genome Sciences, Inc.  
TITLE OF INVENTION: Albumin Fusion Proteins  
FILE REFERENCE: PFS46PCT  
CURRENT APPLICATION NUMBER: US/09/833,245  
CURRENT FILING DATE: 2001-04-12  
PRIOR APPLICATION NUMBER: 60/229,358  
PRIOR FILING DATE: 2000-04-12  
PRIOR APPLICATION NUMBER: 60/256,931  
PRIOR FILING DATE: 2000-12-21  
PRIOR APPLICATION NUMBER: 60/199,384  
PRIOR FILING DATE: 2000-04-25  
NUMBER OF SEQ ID NOS: 2267  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 1474  
LENGTH: 353  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-09-833-245-1474

Query Match 100.0%; Score 228; DB 22; Length 353;  
Best Local Similarity 100.0%; Pred. No. 9.3e-221;  
Matches 228; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MRYLLPSVLLGTAPTYVLAAGVWRLLSAFLEPAFYQALDRLCYQSVWLEFFENYTG 60  
DB 1 MRYLLPSVLLGTAPTYVLAAGVWRLLSAFLEPAFYQALDRLCYQSVWLEFFENYTG 60  
QY 61 VOILLYDGLPKNKENIITYLANHSTVDMIVADILAIQNALGHVRYVLEKGLKPLPGC 120  
DB 61 VOILLYDGLPKNKENIITYLANHSTVDMIVADILAIQNALGHVRYVLEKGLKPLPGC 120  
QY 121 YFAOHGIIYVRSKAKFENEKERNKLOSYVDAGTPMYLVIFPEGTRYNPEOTKVLASQAF 180  
DB 121 YFAOHGIIYVRSKAKFENEKERNKLOSYVDAGTPMYLVIFPEGTRYNPEOTKVLASQAF 180  
QY 181 AAOGLAVLKHVLTLPRIKATHVAFDCMKNYLDALYDVTVYEGKDDG 228  
DB 181 AAOGLAVLKHVLTLPRIKATHVAFDCMKNYLDALYDVTVYEGKDDG 228

RESULT 10  
US-09-833-245-1475  
Sequence 1475, Application US/09833245  
GENERAL INFORMATION:  
APPLICANT: Human Genome Sciences, Inc.  
TITLE OF INVENTION: Albumin Fusion Proteins  
FILE REFERENCE: PFS46PCT  
CURRENT APPLICATION NUMBER: US/09/833,245  
CURRENT FILING DATE: 2001-04-12  
PRIOR APPLICATION NUMBER: 60/229,358  
PRIOR FILING DATE: 2000-04-12  
PRIOR APPLICATION NUMBER: 60/256,931  
PRIOR FILING DATE: 2000-12-21  
PRIOR APPLICATION NUMBER: 60/199,384  
PRIOR FILING DATE: 2000-04-25  
NUMBER OF SEQ ID NOS: 2267



SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 1475  
LENGTH: 353  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-09-853-245-1475

Query Match 100.0% Score 228; DB 22; Length 353;  
Best Local Similarity 100.0%; Pred. No. 9.3e-221;  
Matches 228; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MRYLLPSVVLGTAPTYVLANGVWRLLSAFLPARFYQALDDRLCYCYQSMTLFFFNNTG 60  
DB 1 MRYLLPSVVLGTAPTYVLANGVWRLLSAFLPARFYQALDDRLCYCYQSMTLFFFNNTG 60  
QY 61 VOILLXGDLPRKKNENTYLANHQSIVDWIVADILAIQNALGHVRYVLEKGLKMLPLCYG 120  
DB 61 VOILLXGDLPRKKNENTYLANHQSIVDWIVADILAIQNALGHVRYVLEKGLKMLPLCYG 120  
QY 121 YPAOHGGIVYKRSAPFNEKEMRNKLQSYVDAGTPMYLVTFPEGSTRYNPOTVLSASQAF 180  
DB 121 YPAOHGGIVYKRSAPFNEKEMRNKLQSYVDAGTPMYLVTFPEGSTRYNPOTVLSASQAF 180  
QY 181 AAORGLAVLKHVLPRIKATHVAFCMKNYLDAIVDVYVVEGKDDGG 228  
DB 181 AAORGLAVLKHVLPRIKATHVAFCMKNYLDAIVDVYVVEGKDDGG 228

RESULT 11  
US-09-853-526-4  
Sequence 4, Application US/09853526

GENERAL INFORMATION:

APPLICANT: Cohen, Daniel  
APPLICANT: Blumenfeld, Marta  
APPLICANT: Ilyu, Chumakov  
APPLICANT: Bougueteloret, Lydie  
TITLE OF INVENTION: PROSTATE CANCER GENE  
FILE REFERENCE: GENSET.18CPLCP  
CURRENT APPLICATION NUMBER: US/09/853,526  
CURRENT FILING DATE: 2001-05-11  
PRIOR APPLICATION NUMBER: 09/338,907  
PRIOR FILING DATE: 1999-06-23  
PRIOR APPLICATION NUMBER: 08/996,306  
PRIOR FILING DATE: 1997-12-22  
PRIOR APPLICATION NUMBER: 60/099,658  
PRIOR FILING DATE: 1998-09-09  
PRIOR APPLICATION NUMBER: 09/218,207  
PRIOR FILING DATE: 1998-12-22  
NUMBER OF SEQ ID NOS: 578  
SOFTWARE: Patent.pm  
SEQ ID NO 4

LENGTH: 353

TYPE: PRT

ORGANISM: Homo sapiens

FEATURE:

NAME/KEY: HELIX

LOCATION: 1..33

OTHER INFORMATION: Rao and Argos identification method, potential helix

NAME/KEY: HELIX

LOCATION: 4..20

OTHER INFORMATION: Klein, Kanehisa and Delisi identification method, potential

NAME/KEY: HELIX

LOCATION: 4..24

OTHER INFORMATION: Eisenberg, Schwarz, Komaromy, Wall identification method, F

NAME/KEY: MYRISTATE

LOCATION: 12..16

OTHER INFORMATION: Prosite match

NAME/KEY: HELIX

LOCATION: 50..70

OTHER INFORMATION: Eisenberg, Schwarz, Komaromy, Wall identification method, F

NAME/KEY: HELIX

LOCATION: 50..70

OTHER INFORMATION: Eisenberg, Schwarz, Komaromy, Wall identification method, F

NAME/KEY: HELIX

NAME/KEY: CARBOHYD  
LOCATION: 57..59  
OTHER INFORMATION: Prosite match

NAME/KEY: HELIX

LOCATION: 76..96

OTHER INFORMATION: Eisenberg, Schwarz, Komaromy, Wall identification method,

NAME/KEY: PHOSPHORYLATION

LOCATION: 78

OTHER INFORMATION: potential Tyrosine kinase site, Prosite match

NAME/KEY: PHOSPHORYLATION

LOCATION: 84

OTHER INFORMATION: potential caseine kinase II site, Prosite match

NAME/KEY: SITE

LOCATION: 94..115

OTHER INFORMATION: potential Leucine zipper site, Prosite match

NAME/KEY: MYRISTATE

LOCATION: 119..123

OTHER INFORMATION: potential site, Prosite match

NAME/KEY: PHOSPHORYLATION

LOCATION: 133

OTHER INFORMATION: potential protein kinase C, Prosite match

NAME/KEY: PHOSPHORYLATION

LOCATION: 147

OTHER INFORMATION: potential caseine kinase II site, Prosite match

NAME/KEY: PHOSPHORYLATION

LOCATION: 194

OTHER INFORMATION: potential protein kinase C, Prosite match

NAME/KEY: PHOSPHORYLATION

LOCATION: 215

OTHER INFORMATION: potential Tyrosine kinase site, Prosite match

NAME/KEY: SULFATATION

LOCATION: 221

OTHER INFORMATION: Prosite match

NAME/KEY: PHOSPHORYLATION

LOCATION: 233

OTHER INFORMATION: potential cAMP and cGMP dependant protein kinase site,

NAME/KEY: PHOSPHORYLATION

LOCATION: 235

OTHER INFORMATION: potential caseine kinase II site, Prosite match

NAME/KEY: PHOSPHORYLATION

LOCATION: 306

OTHER INFORMATION: potential protein kinase C, Prosite match

NAME/KEY: HELIX

LOCATION: 310..330

OTHER INFORMATION: Eisenberg, Schwarz, Komaromy, Wall identification method,

NAME/KEY: MYRISTATE

LOCATION: 319..323

OTHER INFORMATION: Prosite match

NAME/KEY: MYRISTATE

LOCATION: 323..327

OTHER INFORMATION: Prosite match

NAME/KEY: AMIDATION

LOCATION: 329

OTHER INFORMATION: Prosite match

NAME/KEY: HELIX

LOCATION: 333..353

OTHER INFORMATION: Eisenberg, Schwarz, Komaromy, Wall identification method,

NAME/KEY: MYRISTATE

LOCATION: 341..345

OTHER INFORMATION: Prosite match

NAME/KEY: PHOSPHORYLATION

LOCATION: 350

OTHER INFORMATION: potential protein kinase C, Prosite match

US-09-853-526-4

Query Match 100.0% Score 228; DB 22; Length 353;

Best Local Similarity 100.0%; Pred. No. 9.3e-221;  
Matches 228; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY 1 MRLPLSVVLGTAPTYVLAAGVWRLISAFPLPARFYQALDDRLCYQSVLFEFENYTG 60
      |||
      1 MRLPLSVVLGTAPTYVLAAGVWRLISAFPLPARFYQALDDRLCYQSVLFEFENYTG 60
Db 1 MRLPLSVVLGTAPTYVLAAGVWRLISAFPLPARFYQALDDRLCYQSVLFEFENYTG 60
      |||
QY 61 VQLLTGDLPKKNENITIIYLANHOSTYDWMIVADILAIROMNLGHVRYLKGKWLPLXGC 120
      |||
      61 VQLLTGDLPKKNENITIIYLANHOSTYDWMIVADILAIROMNLGHVRYLKGKWLPLXGC 120
Db 61 VQLLTGDLPKKNENITIIYLANHOSTYDWMIVADILAIROMNLGHVRYLKGKWLPLXGC 120
      |||
QY 121 YFAHQHGIVYKRSRKAENEKEMRNKLQSYVDAGTPTMYLVIPEEGTRVNEQTKVLSAQAF 180
      |||
      121 YFAHQHGIVYKRSRKAENEKEMRNKLQSYVDAGTPTMYLVIPEEGTRVNEQTKVLSAQAF 180
Db 121 YFAHQHGIVYKRSRKAENEKEMRNKLQSYVDAGTPTMYLVIPEEGTRVNEQTKVLSAQAF 180
      |||
QY 181 AAORGIAVLKHLVLPRIKATVHAFDCMKNYLDAIYDVTVYEGKDDCG 228
      |||
      181 AAORGIAVLKHLVLPRIKATVHAFDCMKNYLDAIYDVTVYEGKDDCG 228
Db 181 AAORGIAVLKHLVLPRIKATVHAFDCMKNYLDAIYDVTVYEGKDDCG 228
      |||

RESULT 12
US-09-901-484A-4
: Sequence 4, Application US/09901484A
: GENERAL INFORMATION:
: APPLICANT: Cohen, Daniel
: APPLICANT: Blumenfeld, Marta
: APPLICANT: Chumakov, Ilya
: APPLICANT: Bougueleret, Lydie
: TITLE OF INVENTION: Prostate Cancer Gene
: FILE REFERENCE: GEN-T11XC3D2
: CURRENT APPLICATION NUMBER: US/09/901,484A
: PRIOR FILING DATE: 2001-07-09
: PRIOR APPLICATION NUMBER: US 08/996,306
: PRIOR FILING DATE: 1997-12-22
: PRIOR APPLICATION NUMBER: US 60/099,658
: PRIOR FILING DATE: 1998-09-09
: PRIOR APPLICATION NUMBER: US 09/218,207
: PRIOR FILING DATE: 1998-12-22
: PRIOR APPLICATION NUMBER: US 09/338,907
: PRIOR FILING DATE: 1999-06-23
: PRIOR APPLICATION NUMBER: US 09/853,526
: PRIOR FILING DATE: 2001-05-11
: NUMBER OF SEQ ID NOS: 578
: SOFTWARE: PatentIn version 3.1
: SEQ ID NO 4
: LENGTH: 353
: TYPE: PRT
: ORGANISM: Homo sapiens
: FEATURE:
: NAME/KEY: HELIX
: LOCATION: (1)..(33)
: OTHER INFORMATION: Rao and Argos identification method, potential helix
: NAME/KEY: HELIX
: LOCATION: (4)..(20)
: OTHER INFORMATION: Klein, Kanehisa and Delisi identification method, potential helix
: NAME/KEY: HELIX
: LOCATION: (4)..(24)
: OTHER INFORMATION: Eisenberg, Schwarz, Komaromy, Wall identification method,
: OTHER INFORMATION: potential helix
: NAME/KEY: LIPID
: LOCATION: (12)..(16)
: OTHER INFORMATION: MYRISTATE, Prosilte match
: NAME/KEY: HELIX
: LOCATION: (50)..(70)
: OTHER INFORMATION: Eisenberg, Schwarz, Komaromy, Wall identification method,
: OTHER INFORMATION: potential helix
: NAME/KEY: CARBOHYD
: LOCATION: (57)..(59)
: OTHER INFORMATION: Prosilte match
: NAME/KEY: HELIX
: LOCATION: (76)..(96)
: OTHER INFORMATION: Eisenberg, Schwarz, Komaromy, Wall identification method,
: OTHER INFORMATION: potential helix
: NAME/KEY: MOD_RES
: LOCATION: (78)..(78)

: OTHER INFORMATION: PHOSPHORYLATION, potential Tyrosine kinase site, Prosilte match
: NAME/KEY: MOD_RES
: LOCATION: (84)..(84)
: OTHER INFORMATION: PHOSPHORYLATION, potential caseine kinase II site, Prosilte mat
: NAME/KEY: SITE
: LOCATION: (94)..(115)
: OTHER INFORMATION: Potential leucine zipper site, Prosilte match
: NAME/KEY: LIPID
: LOCATION: (119)..(123)
: OTHER INFORMATION: MYRISTATE, Prosilte match
: NAME/KEY: MOD_RES
: LOCATION: (133)..(133)
: OTHER INFORMATION: PHOSPHORYLATION, potential protein kinase C, Prosilte match
: NAME/KEY: MOD_RES
: LOCATION: (147)..(147)
: OTHER INFORMATION: PHOSPHORYLATION, potential caseine kinase II, Prosilte match
: NAME/KEY: MOD_RES
: LOCATION: (194)..(194)
: OTHER INFORMATION: PHOSPHORYLATION, potential protein kinase C, Prosilte match
: NAME/KEY: MOD_RES
: LOCATION: (215)..(215)
: OTHER INFORMATION: PHOSPHORYLATION, potential Tyrosine kinase site, Prosilte match
: NAME/KEY: MOD_RES
: LOCATION: (221)..(221)
: OTHER INFORMATION: SULFATATION, Prosilte match
: NAME/KEY: MOD_RES
: LOCATION: (233)..(233)
: OTHER INFORMATION: PHOSPHORYLATION, potential cAMP and cGMP dependant protein kin
: NAME/KEY: MOD_RES
: LOCATION: (235)..(235)
: OTHER INFORMATION: PHOSPHORYLATION, potential caseine kinase II site, Prosilte mat
: NAME/KEY: MOD_RES
: LOCATION: (306)..(306)
: OTHER INFORMATION: PHOSPHORYLATION, potential protein kinase C, Prosilte match
: NAME/KEY: HELIX
: LOCATION: (310)..(330)
: OTHER INFORMATION: Eisenberg, Schwarz, Komaromy, Wall identification method,
: OTHER INFORMATION: potential helix
: NAME/KEY: LIPID
: LOCATION: (319)..(323)
: OTHER INFORMATION: MYRISTATE, Prosilte match
: NAME/KEY: LIPID
: LOCATION: (323)..(327)
: OTHER INFORMATION: MYRISTATE, Prosilte match
: NAME/KEY: MOD_RES
: LOCATION: (329)..(329)
: OTHER INFORMATION: AMIDATION, Prosilte match
: NAME/KEY: HELIX
: LOCATION: (333)..(353)
: OTHER INFORMATION: Eisenberg, Schwarz, Komaromy, Wall identification method,
: OTHER INFORMATION: potential helix
: NAME/KEY: LIPID
: LOCATION: (341)..(345)
: OTHER INFORMATION: MYRISTATE, Prosilte match
: NAME/KEY: MOD_RES
: LOCATION: (350)..(350)
: OTHER INFORMATION: PHOSPHORYLATION, potential protein kinase C, Prosilte match
US-09-901-484A-4

Query Match 100.0%; Score 228; DB 23; Length 353;
Best Local Similarity 100.0%; Pred. No. 9.3e-221;
Matches 228; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MRLPLSVVLGTAPTYVLAAGVWRLISAFPLPARFYQALDDRLCYQSVLFEFENYTG 60
      |||
      1 MRLPLSVVLGTAPTYVLAAGVWRLISAFPLPARFYQALDDRLCYQSVLFEFENYTG 60
Db 1 MRLPLSVVLGTAPTYVLAAGVWRLISAFPLPARFYQALDDRLCYQSVLFEFENYTG 60
      |||
QY 61 VQLLTGDLPKKNENITIIYLANHOSTYDWMIVADILAIROMNLGHVRYLKGKWLPLXGC 120
      |||
      61 VQLLTGDLPKKNENITIIYLANHOSTYDWMIVADILAIROMNLGHVRYLKGKWLPLXGC 120
Db 61 VQLLTGDLPKKNENITIIYLANHOSTYDWMIVADILAIROMNLGHVRYLKGKWLPLXGC 120
      |||
```

Qy	121	YFAOHGGYYVRSKAFNFEKEKRNKLSQSYVAGCPMYLVIPEEGRYRPEDTKYLSSQAF	180
Db	121	YFAOHGGYYVRSKAFNFEKEKRNKLSQSYVAGCPMYLVIPEEGRYRPEDTKYLSSQAF	180
Qy	121	YFAOHGGYYVRSKAFNFEKEKRNKLSQSYVAGCPMYLVIPEEGRYRPEDTKYLSSQAF	180
Db	121	YFAOHGGYYVRSKAFNFEKEKRNKLSQSYVAGCPMYLVIPEEGRYRPEDTKYLSSQAF	180
Qy	181	AAOGLAVLKHVLTLPRIKATHVAFDCKKNTLDAIYDTVYVEGKDGG	228
Db	181	AAOGLAVLKHVLTLPRIKATHVAFDCKKNTLDAIYDTVYVEGKDGG	228

```

RESULT 13
US-09-915-582-59
; Sequence 59, Application US/09915582
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: 17 Human Secreted Proteins
; FILE REFERENCE: P5723p1
; CURRENT APPLICATION NUMBER: US/09/915,582
; PRIOR FILING DATE: 2001-07-27
; PRIOR APPLICATION NUMBER: PCT/US01/01431
; PRIOR FILING DATE: 2001-01-17
; PRIOR APPLICATION NUMBER: 60/179,065
; PRIOR FILING DATE: 2000-01-31
; PRIOR APPLICATION NUMBER: 60/180,628
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: 60/231,968
; PRIOR FILING DATE: 2000-09-12
; NUMBER OF SEQ ID NOS: 97
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 59
; LENGTH: 353
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-09-915-582-59

```

```

RESULT 14
US-09-915-582-79
; Sequence 79, Application US/09915582
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: 17 Human Secreted Proteins
; FILE REFERENCE: PS723P1
; CURRENT APPLICATION NUMBER: US/09/915,582
; CURRENT FILING DATE: 2001-07-27
; PRIOR APPLICATION NUMBER: PCT/US01/01431
; PRIOR FILING DATE: 2001-01-17
; PRIOR APPLICATION NUMBER: 60/179,065
; PRIOR FILING DATE: 2000-01-31
; PRIOR APPLICATION NUMBER: 60/180,628
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: 60/231,968
; PRIOR FILING DATE: 2000-09-12

```

```

: NUMBER OF SEQ ID NOS: 97
: SOFTWARE: PatentIn Ver. 2.0
: SEQ ID NO 79
: LENGTH: 333
: TYPE: PRT
: ORGANISM: Homo sapiens
: OS-09-915-582-79

```

```

1 RESULT 15
2 US-60-099-658-4
3 Sequence 4, Application US/60099658
4 GENERAL INFORMATION:
5 APPLICANT: Cohen, Daniel
6 APPLICANT: Chumakov, Ilya
7 APPLICANT: Blumentfeld, Marta
8 APPLICANT: Bougueleret, Lydie
9 TITLE OF INVENTION: Prostate cancer gene
10 NUMBER OF SEQUENCES: 99
11 CORRESPONDENCE ADDRESS:
12 ADDRESSEE: Knobbe, Martens, Olson & Bear
13 STREET: 501 West Broadway
14 CITY: San Diego
15 STATE: California
16 COUNTRY: USA
17 ZIP: 92101-3505
18 COMPUTER READABLE FORM:
19 MEDIUM TYPE: Floppy Disk
20 COMPUTER: IBM PC compatible
21 OPERATING SYSTEM: win95
22 SOFTWARE: word
23 CURRENT APPLICATION DATA:
24 APPLICATION NUMBER: US/60/099,658
25 FILING DATE:
26 CLASSIFICATION:
27 ATTORNEY/AGENT INFORMATION:
28 NAME: Israelsen, Ned A.
29 REGISTRATION NUMBER: 29,655
30 REFERENCE/DOCKET NUMBER: GENSET .018APR
31 TELECOMMUNICATION INFORMATION:
32 TELEPHONE: (619) 235-8550
33 TELEFAX: (619) 235-0176
34 INFORMATION FOR SEQ ID NO: 4:
35 SEQUENCE CHARACTERISTICS:
36 LENGTH: 353 amino acids
37 TYPE: AMINO ACID
38 STRANDEDNESS: SINGLE
39 TOPOLOGY: LINEAR
40 MOLECULE TYPE: PROTEIN
41 ORIGINAL SOURCE:
42 ORGANISM: Homo sapiens
43 FEATURE:

```

```

? NAME/KEY: potential Transmembrane helix
? LOCATION: 1..33
? IDENTIFICATION METHOD: Rao and Argos method
? FEATURE:
? NAME/KEY: potential Transmembrane helix
? LOCATION: 4..20
? IDENTIFICATION METHOD: Klein, Kanehisa and Delisi method
? FEATURE:
? NAME/KEY: potential Transmembrane helix
? LOCATION: 4..24
? IDENTIFICATION METHOD: Eisenberg,Schwarz,Komarony
? IDENTIFICATION METHOD: and Wall method
? FEATURE:
? NAME/KEY: potential N-myristoylation site
? LOCATION: 12
? IDENTIFICATION METHOD: prosite match
? FEATURE:
? NAME/KEY: potential Transmembrane helix
? LOCATION: 50..70
? IDENTIFICATION METHOD: Eisenberg,Schwarz,Komarony and Wall method
? FEATURE:
? NAME/KEY: potential N-glycosylation site
? LOCATION: 57
? IDENTIFICATION METHOD: prosite match
? FEATURE:
? NAME/KEY: potential Transmembrane helix
? LOCATION: 76..96
? IDENTIFICATION METHOD: Eisenberg,Schwarz,Komarony and Wall method
? FEATURE:
? NAME/KEY: potential Tyrosine kinase phosphorylation site
? LOCATION: 78
? IDENTIFICATION METHOD: prosite match
? FEATURE:
? NAME/KEY: potential Casein kinase II phosphorylation site
? LOCATION: 84
? IDENTIFICATION METHOD: prosite match
? FEATURE:
? NAME/KEY: potential Leucine zipper pattern
? LOCATION: 94..115
? IDENTIFICATION METHOD: prosite match
? FEATURE:
? NAME/KEY: potential N-myristoylation site
? LOCATION: 119
? IDENTIFICATION METHOD: prosite match
? FEATURE:
? NAME/KEY: potential Protein kinase C phosphorylation site
? LOCATION: 133
? IDENTIFICATION METHOD: prosite match
? FEATURE:
? NAME/KEY: potential Casein kinase II phosphorylation site
? LOCATION: 147
? IDENTIFICATION METHOD: prosite match
? FEATURE:
? NAME/KEY: potential Protein kinase C phosphorylation site
? LOCATION: 194
? IDENTIFICATION METHOD: prosite match
? FEATURE:
? NAME/KEY: potential Tyrosine kinase phosphorylation site
? LOCATION: 215
? IDENTIFICATION METHOD: prosite match
? FEATURE:
? NAME/KEY: potential Tyrosine sulfatation site
? LOCATION: 221
? IDENTIFICATION METHOD: prosite match
? FEATURE:
? NAME/KEY: potential cAMP- and cGMP-dependent protein kinase phosphorylation site
? LOCATION: 233
? IDENTIFICATION METHOD: prosite match
? FEATURE:
? NAME/KEY: potential Casein kinase II phosphorylation site
? LOCATION: 235
? IDENTIFICATION METHOD: prosite match
? FEATURE:

```

```

? NAME/KEY: potential Protein kinase C phosphorylation site
? LOCATION: 306
? IDENTIFICATION METHOD: prosite match
? FEATURE:
? NAME/KEY: potential Transmembrane helix
? LOCATION: 310..330
? IDENTIFICATION METHOD: Eisenberg,Schwarz,Komarony and Wall method
? FEATURE:
? NAME/KEY: potential N-myristoylation site
? LOCATION: 319
? IDENTIFICATION METHOD: prosite match
? FEATURE:
? NAME/KEY: potential N-myristoylation site
? LOCATION: 323
? IDENTIFICATION METHOD: prosite match
? FEATURE:
? NAME/KEY: potential Amladation site
? LOCATION: 329
? IDENTIFICATION METHOD: prosite match
? FEATURE:
? NAME/KEY: potential Transmembrane helix
? LOCATION: 333..353
? IDENTIFICATION METHOD: Eisenberg,Schwarz,Komarony and Wall method
? FEATURE:
? NAME/KEY: potential N-myristoylation site
? LOCATION: 341
? IDENTIFICATION METHOD: prosite match
? FEATURE:
? NAME/KEY: potential Protein kinase C phosphorylation site
? LOCATION: 350
? IDENTIFICATION METHOD: prosite match
? US-60-099-658-4

```

```

Query Match 100.0%; Score 228; DB 26; Length 353;
Best Local Similarity 100.0%; Pred No. 9.3e-221;
Matches 228; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MRYLLPSVVLGTAPTYVLAMGWRLSAPLPARYQALDDRLCYQSMVLEFFENYTG 60
Db 1 MRYLLPSVVLGTAPTYVLAMGWRLSAPLPARYQALDDRLCYQSMVLEFFENYTG 60
QY 61 VOILLYGDLEPRKKNENIYLANHOSVDWIVADILAIQNALGHVRYLKEGLKWLPLYGC 120
Db 61 VOILLYGDLEPRKKNENIYLANHOSVDWIVADILAIQNALGHVRYLKEGLKWLPLYGC 120
QY 121 YFAOHGSIYVRSKAKFNEKEMRNKLOSVDAGTPMYLVIPPEGRVYVPEOTKVLASQAF 180
Db 121 YFAOHGSIYVRSKAKFNEKEMRNKLOSVDAGTPMYLVIPPEGRVYVPEOTKVLASQAF 180
QY 181 AAOGLAVLKHVLPRIKATHVAFDCMKNYLDAIYDVTVVYEGKDDG 228
Db 181 AAOGLAVLKHVLPRIKATHVAFDCMKNYLDAIYDVTVVYEGKDDG 228

```

Search completed: August 28, 2002, 11:25:20  
 Job time: 498 sec

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 28, 2002, 11:26:50 ; Search time 83.88 Seconds

(without alignments)  
657.935 Million cell updates/sec

Title: US-09-853-526-70

Perfect score: 228

Sequence: 1 MRYLPSVVLGTAFTYVLA.....NYLDAIVTVVEGKDDGS 228

Scoring table:

Gapop 60.0 , Gapext 60.0

Searched: 747981 seqs, 242050750 residues

Word size : 0

Total number of hits satisfying chosen parameters: 747981

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database : Pending\_Patents\_AA\_New:\*  
1: /cgn2\_6/ptodata/1/paa/PCT\_NEW\_COMB.pep:\*  
2: /cgn2\_6/ptodata/1/paa/US06\_NEW\_COMB.pep:\*  
3: /cgn2\_6/ptodata/1/paa/US07\_NEW\_COMB.pep:\*  
4: /cgn2\_6/ptodata/1/paa/US08\_NEW\_COMB.pep:\*  
5: /cgn2\_6/ptodata/1/paa/US09\_NEW\_COMB.pep:\*  
6: /cgn2\_6/ptodata/1/paa/US10\_NEW\_COMB.pep:\*  
7: /cgn2\_6/ptodata/1/paa/US60\_NEW\_COMB.pep:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	228	100.0	364	US-10-184-648-63	Sequence 63, Appl
2	144	63.2	269	US-10-074-045-47	Sequence 13028, A
3	144	63.2	353	US-09-629-469A-13028	Sequence 52, Appl
4	9	3.9	135	US-10-184-648-52	Sequence 66, Appl
5	9	3.9	157	US-10-184-648-66	Sequence 184, App
6	9	3.9	240	US-09-895-913A-184	Sequence 1538, Ap
7	8	3.5	236	US-10-053-853A-1538	Sequence 5820, Ap
8	8	3.5	262	US-09-791-537-5820	Sequence 110883,
9	8	3.5	459	US-09-791-537-110883	Sequence 142, App
10	8	3.5	483	US-09-791-537-110642	Sequence 150979,
11	7	3.1	44	US-09-826-734A-142	Sequence 31675, A
12	7	3.1	77	US-09-548-936C-17	Sequence 63578, A
13	7	3.1	96	US-09-791-537-150979	Sequence 53250, A
14	7	3.1	117	US-09-791-537-31675	Sequence 9, Appl1
15	7	3.1	126	US-09-791-537-63528	Sequence 51, Appl
16	7	3.1	194	US-09-791-537-53250	Sequence 67, Appl
17	7	3.1	195	US-10-164-966-9	Sequence 86, Appl
18	7	3.1	195	US-10-184-648-51	Sequence 132, App
19	7	3.1	195	US-10-184-648-67	Sequence 57367, A
20	7	3.1	195	US-10-184-648-86	Sequence 99121, A
21	7	3.1	208	PCT-US02-10824-132	Sequence 128069,
22	7	3.1	208	US-09-791-537-57367	Sequence 142901,
23	7	3.1	208	US-09-791-537-99121	Sequence 143774,
24	7	3.1	208	US-09-791-537-128069	
25	7	3.1	208	US-09-791-537-142901	
26	7	3.1	208	US-09-791-537-143774	

## ALIGNMENTS

27	7	3.1	208	US-10-096-327-2	Sequence 2, Appl1
28	7	3.1	208	US-10-179-131-7124	Sequence 7124, Ap
29	7	3.1	208	US-10-189-360-5	Sequence 5, Appl1
30	7	3.1	208	US-10-189-360-7	Sequence 7, Appl1
31	7	3.1	208	US-10-189-360-8	Sequence 8, Appl1
32	7	3.1	208	US-10-138-158-18	Sequence 148811,
33	7	3.1	225	US-09-791-537-148811	Sequence 22428, A
34	7	3.1	254	US-10-155-881-22428	Sequence 751, App
35	7	3.1	262	US-09-791-537-751	Sequence 18654, A
36	7	3.1	263	US-10-135-881-18654	Sequence 1244, Ap
37	7	3.1	269	US-10-053-853A-1244	Sequence 1026, Ap
38	7	3.1	270	US-10-211-364-1026	Sequence 1733, Ap
39	7	3.1	286	US-60-360-039-1733	Sequence 42339, A
40	7	3.1	364	US-09-791-537-42339	Sequence 11833, A
41	7	3.1	400	US-09-791-537-11833	Sequence 116428,
42	7	3.1	409	US-09-791-537-116428	Sequence 94112, A
43	7	3.1	441	US-09-791-537-94112	Sequence 117991,
44	7	3.1	445	US-09-791-537-117991	Sequence 107567,
45	7	3.1	499	US-09-791-537-107567	

RESULT 1  
US-10-184-648-63  
Sequence 63, Application US/10184648  
GENERAL INFORMATION:  
APPLICANT: Meyers, Rachel E.  
APPLICANT: Williamson, Mark  
APPLICANT: Tsai, Fong-Ying  
APPLICANT: Hunter, John J.  
APPLICANT: Macbeth, Kyle J.  
APPLICANT: Rudolph-Owen, Laura A.  
APPLICANT: Leiby, Kevin R.  
APPLICANT: Kapeller-Libermann, Rosana  
APPLICANT: Olandt, Peter J.  
TITLE OF INVENTION: NOVEL HUMAN TRANSFERASE FAMILY MEMBERS AND USES THEREOF  
FILE REFERENCE: 10448-192001  
CURRENT APPLICATION NUMBER: US/10/184, 648  
CURRENT FILING DATE: 2002-06-27  
PRIOR APPLICATION NUMBER: US 09/815, 028  
PRIOR FILING DATE: 2001-03-22  
PRIOR APPLICATION NUMBER: PCT/US01/09358  
PRIOR FILING DATE: 2001-03-22  
PRIOR APPLICATION NUMBER: US 60/191, 964  
PRIOR FILING DATE: 2000-03-24  
PRIOR APPLICATION NUMBER: US 09/801, 220  
PRIOR FILING DATE: 2001-03-07  
PRIOR APPLICATION NUMBER: PCT/US01/07269  
PRIOR FILING DATE: 2001-03-07  
PRIOR APPLICATION NUMBER: US 60/187, 456  
PRIOR FILING DATE: 2000-03-07  
PRIOR APPLICATION NUMBER: US 09/816, 714  
PRIOR FILING DATE: 2001-03-23  
PRIOR APPLICATION NUMBER: PCT/US01/09468  
PRIOR FILING DATE: 2001-03-23  
PRIOR APPLICATION NUMBER: US 60/191, 865  
PRIOR FILING DATE: 2000-03-24  
PRIOR APPLICATION NUMBER: US 09/844, 948  
PRIOR FILING DATE: 2001-04-27  
PRIOR APPLICATION NUMBER: PCT/US01/13805  
PRIOR FILING DATE: 2001-04-27  
PRIOR APPLICATION NUMBER: US 60/200, 604  
PRIOR FILING DATE: 2000-04-28  
PRIOR APPLICATION NUMBER: US 09/861, 164  
PRIOR FILING DATE: 2001-05-18  
PRIOR APPLICATION NUMBER: PCT/US01/16292  
PRIOR FILING DATE: 2001-05-18  
PRIOR APPLICATION NUMBER: US 60/205, 408  
PRIOR FILING DATE: 2000-05-19  
PRIOR APPLICATION NUMBER: US 09/883, 060  
PRIOR FILING DATE: 2001-06-15

```

; PRIOR APPLICATION NUMBER: PCT/US01/19138
; PRIOR FILING DATE: 2001-06-15
; PRIOR APPLICATION NUMBER: US 60/212,079
; PRIOR FILING DATE: 2000-06-15
; PRIOR APPLICATION NUMBER: US 09/962,678
; PRIOR FILING DATE: 2001-09-25
; PRIOR APPLICATION NUMBER: PCT/US01/29963
; PRIOR FILING DATE: 2001-09-25
; PRIOR APPLICATION NUMBER: US 60/235,044
; PRIOR FILING DATE: 2000-09-25
; PRIOR APPLICATION NUMBER: US 09/973,457
; PRIOR FILING DATE: 2001-10-09
; PRIOR APPLICATION NUMBER: US 60/238,849
; PRIOR FILING DATE: 2000-10-06
; PRIOR APPLICATION NUMBER: US 10/072,285
; PRIOR FILING DATE: 2002-02-08
; PRIOR APPLICATION NUMBER: PCT/US02/03736
; PRIOR FILING DATE: 2002-02-08
; PRIOR APPLICATION NUMBER: US 60/267,494
; PRIOR FILING DATE: 2001-02-08
; PRIOR APPLICATION NUMBER: US 09/817,910
; PRIOR FILING DATE: 2001-03-26
; PRIOR APPLICATION NUMBER: PCT/US01/09633
; PRIOR FILING DATE: 2001-03-26
; PRIOR APPLICATION NUMBER: US 60/192,092
; PRIOR FILING DATE: 2000-03-24
; PRIOR APPLICATION NUMBER: US 09/842,528
; PRIOR FILING DATE: 2001-04-25
; PRIOR APPLICATION NUMBER: PCT/US01/40607
; PRIOR FILING DATE: 2001-04-25
; PRIOR APPLICATION NUMBER: US 60/199,500
; PRIOR FILING DATE: 2000-04-25
; PRIOR APPLICATION NUMBER: US 09/882,836
; PRIOR FILING DATE: 2001-06-15
; PRIOR APPLICATION NUMBER: PCT/US01/19543
; PRIOR FILING DATE: 2001-06-15
; PRIOR APPLICATION NUMBER: US 60/211,730
; PRIOR FILING DATE: 2000-06-15
; PRIOR APPLICATION NUMBER: US 09/882,872
; PRIOR FILING DATE: 2001-06-15
; PRIOR APPLICATION NUMBER: PCT/US01/19153
; PRIOR FILING DATE: 2001-06-15
; PRIOR APPLICATION NUMBER: US 60/212,077
; PRIOR FILING DATE: 2000-06-15
; NUMBER OF SEQ ID NOS: 90
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 63
; LENGTH: 364
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-184-648-63
```

```
Query Match          100.0%; Score 228; DB 6; Length 364;
Best Local Similarity 100.0%; Pred. No. 2.9e-225;
Matches 228; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```

QY 1 MRYLLPSVLLGTAFTYVLAAGVWRLSAPFLPARFYQALDDRLCYQOSMWLFFFEENTG 60
    |||||
DB 12 MRYLLPSVLLGTAFTYVLAAGVWRLSAPFLPARFYQALDDRLCYQOSMWLFFFEENTG 71
    |||||
QY 61 VQILLYDDELRKKNENIYLANHCSIVDMIVADILAIROMALGHVRYVLEKGLKWLPLYGC 120
    |||||
DB 72 VQILLYDDELRKKNENIYLANHCSIVDMIVADILAIROMALGHVRYVLEKGLKWLPLYGC 131
    |||||
QY 121 YFAQHGIIYVYKRSKAFNEKEKRNKLSQSYVDAQTPMYLVIPEEGRTYRNPEDQTKVLSASQAF 180
    |||||
DB 132 YFAQHGIIYVYKRSKAFNEKEKRNKLSQSYVDAQTPMYLVIPEEGRTYRNPEDQTKVLSASQAF 191
    |||||
QY 181 AAGRGVLAVLKHVLTPIKATHVAFDCKMKNYLDATYDVTVYVEGKDDGG 228
    |||||
DB 192 AAGRGVLAVLKHVLTPIKATHVAFDCKMKNYLDATYDVTVYVEGKDDGG 239
    |||||
```

```

RESULT 2
US-10-074-045-47
; Sequence 47, Application US/10074045
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PT221C1
; CURRENT APPLICATION NUMBER: US/10/074,045
; CURRENT FILING DATE: 2002-02-14
; Prior Application removed - See file wrapper or Palm
; NUMBER OF SEQ ID NOS: 74
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 47
; LENGTH: 269
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-074-045-47
```

```
Query Match          63.2%; Score 144; DB 6; Length 269;
Best Local Similarity 100.0%; Pred. No. 3.1e-139;
Matches 144; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```

QY 85 TVDMIVADILAIROMALGHVRYVLEKGLKWLPLYGCTFPAQHGIIYVYKRSKAFNEKEKRNK 144
    |||||
DB 1 TVDMIVADILAIROMALGHVRYVLEKGLKWLPLYGCTFPAQHGIIYVYKRSKAFNEKEKRNK 60
    |||||
QY 145 LOSYVDAGTPMYLVIPEEGRTYRNPEDQTKVLSASQAFPAQGLAVLKHVLTPIKATHVAF 204
    |||||
DB 61 LOSYVDAGTPMYLVIPEEGRTYRNPEDQTKVLSASQAFPAQGLAVLKHVLTPIKATHVAF 120
    |||||
QY 205 DCKMKNYLDATYDVTVYVEGKDDGG 228
    |||||
DB 121 DCKMKNYLDATYDVTVYVEGKDDGG 144
    |||||
```

```

RESULT 3
US-09-629-469A-13028
; Sequence 13028, Application US/09629469A
; GENERAL INFORMATION:
; APPLICANT: OTA, TOSHIO
; APPLICANT: ISOGAI, TAKAO
; APPLICANT: NISHIKAWA, TETSUO
; APPLICANT: HAYASHI, KOJI
; APPLICANT: SAITO, KAORI
; APPLICANT: YAMAMOTO, JUNICHI
; APPLICANT: ISHII, SHIZUKO
; APPLICANT: SUGIYAMA, TOMOYASU
; APPLICANT: WAKAMATSU, AI
; APPLICANT: NAGAI, KEIICHI
; APPLICANT: OTSUKI, TETSUJI
; TITLE OF INVENTION: PRIMERS FOR SYNTHESIZING FULL-LENGTH CDNA AND THEIR USE
; FILE REFERENCE: 084335/0123
; CURRENT APPLICATION NUMBER: US/09/629,469A
; CURRENT FILING DATE: 2000-07-28
; PRIOR APPLICATION NUMBER: JP 1999-248036
; PRIOR FILING DATE: 1999-07-29
; PRIOR APPLICATION NUMBER: JP 1999-300253
; PRIOR FILING DATE: 1999-08-27
; PRIOR APPLICATION NUMBER: JP 2000-118776
; PRIOR FILING DATE: 2000-01-11
; PRIOR APPLICATION NUMBER: JP 2000-183767
; PRIOR FILING DATE: 2000-05-02
; PRIOR APPLICATION NUMBER: JP 2000-241899
; PRIOR FILING DATE: 2000-06-09
; PRIOR APPLICATION NUMBER: 60/159,590
; PRIOR FILING DATE: 1999-10-18
; PRIOR APPLICATION NUMBER: 60/183,322
; PRIOR FILING DATE: 2000-02-17
; NUMBER OF SEQ ID NOS: 19025
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 13028
```

LENGTH: 353  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-09-629-469A-13028

Query Match 63.2%; Score 144; DB 5; Length 353;  
Best Local Similarity 100.0%; Pred. No. 4e-139;  
Matches 144; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 MRYLLPSVVLGTAPTYVLAMGVWRLLSAFLPAREYQALDDRLCYVQSMVLFEEFNNTG 60  
DB 1 MRYLLPSVVLGTAPTYVLAMGVWRLLSAFLPAREYQALDDRLCYVQSMVLFEEFNNTG 60  
OY 61 VOILLYGDLPKKNENIYLANHOSYVDVADILAIROMALGHVRYLKEGLKMLPLXGC 120  
DB 61 VOILLYGDLPKKNENIYLANHOSYVDVADILAIROMALGHVRYLKEGLKMLPLXGC 120  
OY 121 YFAOHGIYVKRSKAKFNEKEMRNK 144  
DB 121 YFAOHGIYVKRSKAKFNEKEMRNK 144

RESULT 4  
US-10-184-648-52  
Sequence 52, Application US/10184648

GENERAL INFORMATION:  
APPLICANT: Meyers, Rachel E.  
APPLICANT: Williamson, Mark  
APPLICANT: Tsal, Fong-Ying  
APPLICANT: Hunter, John J.  
APPLICANT: Macbeth, Kyle J.  
APPLICANT: Rudolph-Owen, Laura A.  
APPLICANT: Leiby, Kevin R.  
APPLICANT: Kapeller-Libermann, Rosana  
APPLICANT: Olandt, Peter J.  
TITLE OF INVENTION: NOVEL HUMAN TRANSFERASE FAMILY MEMBERS AND USES THEREOF  
FILE REFERENCE: 10448-192001  
CURRENT APPLICATION NUMBER: US/10/184,648  
CURRENT FILING DATE: 2002-06-27  
PRIOR APPLICATION NUMBER: US 09/815,028  
PRIOR FILING DATE: 2001-03-22  
PRIOR APPLICATION NUMBER: PCT/US01/09358  
PRIOR FILING DATE: 2001-03-22  
PRIOR APPLICATION NUMBER: US 60/191,964  
PRIOR FILING DATE: 2000-03-24  
PRIOR APPLICATION NUMBER: US 09/801,220  
PRIOR FILING DATE: 2001-03-07  
PRIOR APPLICATION NUMBER: PCT/US01/07269  
PRIOR FILING DATE: 2001-03-07  
PRIOR APPLICATION NUMBER: US 60/187,456  
PRIOR FILING DATE: 2000-03-07  
PRIOR APPLICATION NUMBER: US 09/816,714  
PRIOR FILING DATE: 2001-03-23  
PRIOR APPLICATION NUMBER: PCT/US01/09468  
PRIOR FILING DATE: 2001-03-23  
PRIOR APPLICATION NUMBER: US 60/191,865  
PRIOR FILING DATE: 2000-03-24  
PRIOR APPLICATION NUMBER: US 09/844,948  
PRIOR FILING DATE: 2001-04-27  
PRIOR APPLICATION NUMBER: PCT/US01/13805  
PRIOR FILING DATE: 2001-04-27  
PRIOR APPLICATION NUMBER: US 60/200,604  
PRIOR FILING DATE: 2000-04-28  
PRIOR APPLICATION NUMBER: US 09/861,164  
PRIOR FILING DATE: 2001-05-18  
PRIOR APPLICATION NUMBER: PCT/US01/16292  
PRIOR FILING DATE: 2001-05-18  
PRIOR APPLICATION NUMBER: US 60/205,408  
PRIOR FILING DATE: 2000-05-19  
PRIOR APPLICATION NUMBER: US 09/883,060  
PRIOR FILING DATE: 2001-06-15  
PRIOR APPLICATION NUMBER: PCT/US01/19138

PRIOR FILING DATE: 2001-06-15  
PRIOR APPLICATION NUMBER: US 60/212,079  
PRIOR FILING DATE: 2000-06-15  
PRIOR APPLICATION NUMBER: US 09/962,678  
PRIOR FILING DATE: 2001-09-25  
PRIOR APPLICATION NUMBER: PCT/US01/29963  
PRIOR FILING DATE: 2001-09-25  
PRIOR APPLICATION NUMBER: US 60/235,044  
PRIOR FILING DATE: 2000-09-25  
PRIOR APPLICATION NUMBER: US 09/973,457  
PRIOR FILING DATE: 2001-10-09  
PRIOR APPLICATION NUMBER: US 60/238,849  
PRIOR FILING DATE: 2000-10-06  
PRIOR APPLICATION NUMBER: US 10/072,285  
PRIOR FILING DATE: 2002-02-08  
PRIOR APPLICATION NUMBER: PCT/US02/03736  
PRIOR FILING DATE: 2002-02-08  
PRIOR APPLICATION NUMBER: US 60/267,494  
PRIOR FILING DATE: 2001-02-08  
PRIOR APPLICATION NUMBER: US 09/817,910  
PRIOR FILING DATE: 2001-03-26  
PRIOR APPLICATION NUMBER: PCT/US01/09633  
PRIOR FILING DATE: 2001-03-26  
PRIOR APPLICATION NUMBER: US 60/192,092  
PRIOR FILING DATE: 2000-03-24  
PRIOR APPLICATION NUMBER: US 09/842,528  
PRIOR FILING DATE: 2001-04-25  
PRIOR APPLICATION NUMBER: PCT/US01/40607  
PRIOR FILING DATE: 2001-04-25  
PRIOR APPLICATION NUMBER: US 60/199,500  
PRIOR FILING DATE: 2000-04-25  
PRIOR APPLICATION NUMBER: US 09/882,836  
PRIOR FILING DATE: 2001-06-15  
PRIOR APPLICATION NUMBER: PCT/US01/19543  
PRIOR FILING DATE: 2001-06-15  
PRIOR APPLICATION NUMBER: US 60/211,730  
PRIOR FILING DATE: 2000-06-15  
PRIOR APPLICATION NUMBER: US 09/882,872  
PRIOR FILING DATE: 2001-06-15  
PRIOR APPLICATION NUMBER: PCT/US01/19153  
PRIOR FILING DATE: 2001-06-15  
PRIOR APPLICATION NUMBER: US 60/212,077  
PRIOR FILING DATE: 2000-06-15  
NUMBER OF SEQ ID NOS: 90  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 52  
LENGTH: 135  
TYPE: PRT  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: consensus sequence  
US-10-184-648-52

Query Match 3.9%; Score 9; DB 6; Length 135;  
Best Local Similarity 100.0%; Pred. No. 0.48;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 157 LVIFPEGTR 165  
DB 85 LVIFPEGTR 93

RESULT 5  
US-10-184-648-66  
Sequence 66, Application US/10184648  
GENERAL INFORMATION:  
APPLICANT: Meyers, Rachel E.  
APPLICANT: Williamson, Mark  
APPLICANT: Tsal, Fong-Ying  
APPLICANT: Hunter, John J.  
APPLICANT: Macbeth, Kyle J.  
APPLICANT: Rudolph-Owen, Laura A.

APPLICANT: Leiby, Kevin R.  
 APPLICANT: Kapeller-Liebermann, Rosana  
 APPLICANT: Olandt, Peter J.  
 TITLE OF INVENTION: NOVEL HUMAN TRANSFERASE FAMILY MEMBERS AND USES THEREOF  
 FILE REFERENCE: 10448-192001  
 CURRENT APPLICATION NUMBER: US/10/184,648  
 CURRENT FILING DATE: 2002-06-27  
 PRIOR APPLICATION NUMBER: US 09/815,028  
 PRIOR FILING DATE: 2001-03-22  
 PRIOR APPLICATION NUMBER: PCT/US01/09358  
 PRIOR FILING DATE: 2001-03-22  
 PRIOR APPLICATION NUMBER: US 60/191,964  
 PRIOR FILING DATE: 2000-03-24  
 PRIOR APPLICATION NUMBER: US 09/801,220  
 PRIOR FILING DATE: 2001-03-07  
 PRIOR APPLICATION NUMBER: PCT/US01/07269  
 PRIOR FILING DATE: 2001-03-07  
 PRIOR APPLICATION NUMBER: US 60/187,456  
 PRIOR FILING DATE: 2000-03-07  
 PRIOR APPLICATION NUMBER: US 09/816,714  
 PRIOR FILING DATE: 2001-03-23  
 PRIOR APPLICATION NUMBER: PCT/US01/09468  
 PRIOR FILING DATE: 2001-03-23  
 PRIOR APPLICATION NUMBER: US 60/191,865  
 PRIOR FILING DATE: 2000-03-24  
 PRIOR APPLICATION NUMBER: US 09/844,948  
 PRIOR FILING DATE: 2001-04-27  
 PRIOR APPLICATION NUMBER: PCT/US01/13805  
 PRIOR FILING DATE: 2001-04-27  
 PRIOR APPLICATION NUMBER: US 60/200,604  
 PRIOR FILING DATE: 2000-04-28  
 PRIOR APPLICATION NUMBER: US 09/861,164  
 PRIOR FILING DATE: 2001-05-18  
 PRIOR APPLICATION NUMBER: PCT/US01/16292  
 PRIOR FILING DATE: 2001-05-18  
 PRIOR APPLICATION NUMBER: US 60/205,408  
 PRIOR FILING DATE: 2000-05-19  
 PRIOR APPLICATION NUMBER: US 09/883,060  
 PRIOR FILING DATE: 2001-06-15  
 PRIOR APPLICATION NUMBER: PCT/US01/19138  
 PRIOR FILING DATE: 2001-06-15  
 PRIOR APPLICATION NUMBER: US 60/212,079  
 PRIOR FILING DATE: 2000-06-15  
 PRIOR APPLICATION NUMBER: US 09/962,678  
 PRIOR FILING DATE: 2001-09-25  
 PRIOR APPLICATION NUMBER: PCT/US01/29963  
 PRIOR FILING DATE: 2001-09-25  
 PRIOR APPLICATION NUMBER: US 60/235,044  
 PRIOR FILING DATE: 2000-09-25  
 PRIOR APPLICATION NUMBER: US 09/973,457  
 PRIOR FILING DATE: 2001-10-09  
 PRIOR APPLICATION NUMBER: US 60/238,849  
 PRIOR FILING DATE: 2000-10-06  
 PRIOR APPLICATION NUMBER: US 10/072,285  
 PRIOR FILING DATE: 2002-02-08  
 PRIOR APPLICATION NUMBER: PCT/US02/03736  
 PRIOR FILING DATE: 2002-02-08  
 PRIOR APPLICATION NUMBER: US 60/267,494  
 PRIOR FILING DATE: 2001-02-08  
 PRIOR APPLICATION NUMBER: US 09/817,910  
 PRIOR FILING DATE: 2001-03-26  
 PRIOR APPLICATION NUMBER: PCT/US01/09633  
 PRIOR FILING DATE: 2001-03-26  
 PRIOR APPLICATION NUMBER: US 60/192,092  
 PRIOR FILING DATE: 2000-03-24  
 PRIOR APPLICATION NUMBER: US 03/842,528  
 PRIOR FILING DATE: 2001-04-25  
 PRIOR APPLICATION NUMBER: PCT/US01/40607  
 PRIOR FILING DATE: 2001-04-25  
 PRIOR APPLICATION NUMBER: US 60/199,500  
 PRIOR FILING DATE: 2000-04-25  
 PRIOR APPLICATION NUMBER: US 09/882,836  
 PRIOR FILING DATE: 2001-06-15

PRIOR APPLICATION NUMBER: PCT/US01/19543  
 PRIOR FILING DATE: 2001-06-15  
 PRIOR APPLICATION NUMBER: US 60/211,730  
 PRIOR FILING DATE: 2000-06-15  
 PRIOR APPLICATION NUMBER: US 09/882,872  
 PRIOR FILING DATE: 2001-06-15  
 PRIOR APPLICATION NUMBER: PCT/US01/19153  
 PRIOR FILING DATE: 2001-06-15  
 PRIOR APPLICATION NUMBER: US 60/212,077  
 PRIOR FILING DATE: 2000-06-15  
 NUMBER OF SEQ ID NOS: 90  
 SOFTWARE: FastSeq for Windows Version 4.0  
 SEQ ID NO 66  
 LENGTH: 157  
 TYPE: PRT  
 ORGANISM: Artificial Sequence  
 FEATURE:  
 OTHER INFORMATION: consensus sequence  
 US-10-184-648-66

Query Match 3.9%; Score 9; DB 6; Length 157;  
 Best Local Similarity 100.0%; Pred. No. 0.55;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 211 LDAIYDVT 219  
 Db 4 LDAIYDVT 12

RESULT 6  
 US-09-895-913A-184  
 Sequence 184, Application US/09895913A  
 GENERAL INFORMATION:  
 APPLICANT: Kleenhouse, Harold  
 APPLICANT: Al-Garawi, Amal  
 APPLICANT: Miller, Charles  
 APPLICANT: Tomb, Jean Francois  
 TITLE OF INVENTION: Identification of Polynucleotides  
 TITLE OF INVENTION: Encoding Novel Helicobacter Polypeptides in the Helicobacter  
 TITLE OF INVENTION: Genome  
 FILE REFERENCE: 06132/043002  
 CURRENT APPLICATION NUMBER: US/09/895,913A  
 CURRENT FILING DATE: 2001-06-29  
 PRIOR APPLICATION NUMBER: US 08/881,227  
 PRIOR FILING DATE: 1997-06-24  
 NUMBER OF SEQ ID NOS: 368  
 SOFTWARE: FastSeq for Windows Version 4.0  
 SEQ ID NO 184  
 LENGTH: 240  
 TYPE: PRT  
 ORGANISM: Helicobacter pylori  
 US-09-895-913A-184

Query Match 3.9%; Score 9; DB 5; Length 240;  
 Best Local Similarity 100.0%; Pred. No. 0.82;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 157 LVIFPESTR 165  
 Db 142 LVIFPESTR 150

RESULT 7  
 US-10-053-853A-1538  
 Sequence 1538, Application US/10053853A  
 GENERAL INFORMATION:  
 APPLICANT: HAYASHI, Hideo  
 APPLICANT: SHINAGAWA, Hideo  
 APPLICANT: MAKINO, Kozo  
 APPLICANT: HAYASHI, Tetsuya



```
; APPLICANT: OHNISHI, Makoto
; APPLICANT: HATTORI, Masahira
; APPLICANT: KUROKAWA, Ken
; TITLE OF INVENTION: Polynucleotide molecules and polypeptides specific to Enterohemorrhagic E. coli O157:H7
; FILE REFERENCE: 2002-0060A/MWC/01704
; CURRENT APPLICATION NUMBER: US/10/053,853A
; PRIOR FILING DATE: 2002-05-28
; PRIOR APPLICATION NUMBER: JP2001-112010
; NUMBER OF SEQ ID NOS: 1866
; SEQ ID NO 1538
; LENGTH: 236
; TYPE: PRT
; ORGANISM: Escherichia coli O157:H7
US-10-053-853A-1538

Query Match
Best Local Similarity 3.5%; Score 8; DB 6; Length 236;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 193 LPPRIKAT 200
Db 227 LPPRIKAT 234

RESULT 8
US-09-791-537-5820
; Sequence 5820, Application US/09791537
; GENERAL INFORMATION:
; APPLICANT: Bionomix, Inc.
; APPLICANT: Debe, Derek
; APPLICANT: Danzer, Joseph
; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMBERS
; FILE REFERENCE: 261/210
; CURRENT APPLICATION NUMBER: US/09/791,537
; CURRENT FILING DATE: 2001-02-22
; NUMBER OF SEQ ID NOS: 153055
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 5820
; LENGTH: 262
; TYPE: PRT
; ORGANISM: Chlamydia trachomatis
US-09-791-537-5820

Query Match
Best Local Similarity 3.5%; Score 8; DB 5; Length 262;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 191 HVLTPRIK 198
Db 202 HVLTPRIK 209

RESULT 9
US-09-791-537-110883
; Sequence 110883, Application US/09791537
; GENERAL INFORMATION:
; APPLICANT: Bionomix, Inc.
; APPLICANT: Debe, Derek
; APPLICANT: Danzer, Joseph
; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMBERS
; FILE REFERENCE: 261/210
; CURRENT APPLICATION NUMBER: US/09/791,537
; CURRENT FILING DATE: 2001-02-22
; NUMBER OF SEQ ID NOS: 153055
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 110883
; LENGTH: 459
```

```
; TYPE: PRT
; ORGANISM: Bacillus subtilis
US-09-791-537-110883

Query Match
Best Local Similarity 3.5%; Score 8; DB 5; Length 459;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 167 NPEQTKVL 174
Db 292 NPEQTKVL 299

RESULT 10
US-09-791-537-110642
; Sequence 110642, Application US/09791537
; GENERAL INFORMATION:
; APPLICANT: Bionomix, Inc.
; APPLICANT: Debe, Derek
; APPLICANT: Danzer, Joseph
; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMBERS
; FILE REFERENCE: 261/210
; CURRENT APPLICATION NUMBER: US/09/791,537
; CURRENT FILING DATE: 2001-02-22
; NUMBER OF SEQ ID NOS: 153055
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 110642
; LENGTH: 483
; TYPE: PRT
; ORGANISM: Pisum sativum
US-09-791-537-110642

Query Match
Best Local Similarity 3.5%; Score 8; DB 5; Length 483;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 151 AGTPMYLY 158
Db 184 AGTPMYLY 191

RESULT 11
US-09-826-734A-142
; Sequence 142, Application US/09826734A
; GENERAL INFORMATION:
; APPLICANT: Fernandes, Elma R
; APPLICANT: Mishra, Vishnu S
; APPLICANT: Leach, Martin D
; APPLICANT: Shimkets, Richard A
; APPLICANT: Zerhusen, Bryan D
; APPLICANT: Kekuda, Ramesh
; TITLE OF INVENTION: Novel Polynucleotides And Polypeptides Encoded Thereby
; FILE REFERENCE: 15966-754
; CURRENT APPLICATION NUMBER: US/09/826,734A
; CURRENT FILING DATE: 2001-04-05
; PRIOR APPLICATION NUMBER: 60/195,576
; PRIOR FILING DATE: 2000-04-26
; NUMBER OF SEQ ID NOS: 264
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 142
; LENGTH: 44
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-826-734A-142

Query Match
Best Local Similarity 3.1%; Score 7; DB 5; Length 44;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

OY 159 IPEPCTR 165  
|||||||  
Db 3 IPEPCTR 9

RESULT 12  
US-09-548-936C-17  
; Sequence 17, Application US/09548936C  
; GENERAL INFORMATION:  
; APPLICANT: Gomez, Leonel Jorge  
; APPLICANT: Sarras, Jan  
; APPLICANT: Claesson-Weish, Lena  
; APPLICANT: Heidlin, Carl-Jensrik  
; TITLE OF INVENTION: PTPL BINDING AGENTS  
; FILE REFERENCE: L0461/7084  
; CURRENT APPLICATION NUMBER: US/09/548,936C  
; CURRENT FILING DATE: 2000-04-13  
; PRIOR APPLICATION NUMBER: US 09/100,804  
; PRIOR FILING DATE: 1998-06-19  
; PRIOR APPLICATION NUMBER: US 08/596,291  
; PRIOR FILING DATE: 1994-09-01  
; PRIOR APPLICATION NUMBER: US 08/115,573  
; PRIOR FILING DATE: 1993-09-01  
; NUMBER OF SEQ ID NOS: 35  
; SOFTWARE: PatentIn Version 3.0  
; SEQ ID NO 17  
; LENGTH: 77  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-548-936C-17

Query Match 3.1%; Score 7; DB 5; Length 77;  
Best Local Similarity 100.0%; Pred. No. 31;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 125 HGGIYVK 131  
|||||||  
Db 20 HGGIYVK 26

RESULT 13  
US-09-791-537-150979  
; Sequence 150979, Application US/09791537  
; GENERAL INFORMATION:  
; APPLICANT: Bionomix, Inc.  
; APPLICANT: Danzer, Derek  
; APPLICANT: Danzer, Joseph  
; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMB  
; FILE REFERENCE: 261/210  
; CURRENT APPLICATION NUMBER: US/09/791,537  
; CURRENT FILING DATE: 2001-02-22  
; NUMBER OF SEQ ID NOS: 153055  
; SOFTWARE: PatentIn Version 3.0  
; SEQ ID NO 150979  
; LENGTH: 96  
; TYPE: PRT  
; ORGANISM: pdb 3PDZA  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: (1)..(1)  
; OTHER INFORMATION: X is an unknown amino acid  
US-09-791-537-150979

Query Match 3.1%; Score 7; DB 5; Length 96;  
Best Local Similarity 100.0%; Pred. No. 39;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 125 HGGIYVK 131  
|||||||  
Db 32 HGGIYVK 38

RESULT 14  
US-09-791-537-31675  
; Sequence 31675, Application US/09791537  
; GENERAL INFORMATION:  
; APPLICANT: Bionomix, Inc.  
; APPLICANT: Debe, Derek  
; APPLICANT: Danzer, Joseph  
; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY ME  
; FILE REFERENCE: 261/210  
; CURRENT APPLICATION NUMBER: US/09/791,537  
; CURRENT FILING DATE: 2001-02-22  
; NUMBER OF SEQ ID NOS: 153055  
; SOFTWARE: PatentIn Version 3.0  
; SEQ ID NO 31675  
; LENGTH: 117  
; TYPE: PRT  
; ORGANISM: Mus musculus  
US-09-791-537-31675

Query Match 3.1%; Score 7; DB 5; Length 117;  
Best Local Similarity 100.0%; Pred. No. 47;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 125 HGGIYVK 131  
|||||||  
Db 43 HGGIYVK 49

RESULT 15  
US-09-791-537-63528  
; Sequence 63528, Application US/09791537  
; GENERAL INFORMATION:  
; APPLICANT: Bionomix, Inc.  
; APPLICANT: Debe, Derek  
; APPLICANT: Danzer, Joseph  
; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY ME  
; FILE REFERENCE: 261/210  
; CURRENT APPLICATION NUMBER: US/09/791,537  
; CURRENT FILING DATE: 2001-02-22  
; NUMBER OF SEQ ID NOS: 153055  
; SOFTWARE: PatentIn Version 3.0  
; SEQ ID NO 63528  
; LENGTH: 126  
; TYPE: PRT  
; ORGANISM: Mus musculus  
US-09-791-537-63528

Query Match 3.1%; Score 7; DB 5; Length 126;  
Best Local Similarity 100.0%; Pred. No. 50;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 125 HGGIYVK 131  
|||||||  
Db 52 HGGIYVK 58

Search completed: August 28, 2002, 11:26:51  
Job time: 564 sec

8

**THIS PAGE BLANK (USPTO)**

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 28, 2002, 11:27:18 ; Search time 21.78 Seconds  
(without alignments)  
647.104 Million cell updates/sec

Title: US-09-853-526-5

Perfect score: 364  
Sequence: 1 MLSTLVHTFSMRYLPSV.....YVNWIVYTLGCLWYIKA 364

Scoring table:

OLIGO  
Gapop 60.0 , Gapext 60.0

Searched: 105224 seqs, 38719550 residues

Word size : 0

Total number of hits satisfying chosen parameters: 105224

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database : SwissProt\_40:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	353	97.0	353	P1CE_HUMAN	Q9N022 homo sapien
2	32	8.8	354	P1CE_MOUSE	Q9D188 mus musculu
3	9	2.5	237	P1SC_HELPY	O9Z1N8 heliobacte
4	9	2.5	240	P1SC_HELPY	O25003 heliobacte
5	8	2.2	83	V187_BPT3	P10302 bacterioph
6	8	2.2	83	V187_BPT3	P03788 bacterioph
7	8	2.2	396	YD18_YEAST	Q12185 saccharomyc
8	2.2	611	1	HSCA_BUCAL	P57660 buchnera ap
9	7	1.9	57	YUDO_ECOLI	P58038 escherichia
10	7	1.9	194	PURL_IACCA	P35853 lactobacill
11	7	1.9	208	HBCF_CERAE	O09118 ceratophthec
12	7	1.9	208	HBCF_HUMAN	O99075 homo sapien
13	7	1.9	208	HBCF_PIG	O01580 sus scrofa
14	7	1.9	250	HXB4_MOUSE	Q06175 rattus norv
15	7	1.9	250	HXB4_MOUSE	P10284 mus musculu
16	7	1.9	250	HXB4_MOUSE	O59188 borrelia bu
17	7	1.9	250	HXB4_MOUSE	Q23449 caenorhabdi
18	7	1.9	251	HXB4_HUMAN	P17483 homo sapien
19	7	1.9	286	PROC_YEAST	P32263 saccharomyc
20	7	1.9	308	PLSC_COCCNU	Q24270 cocco nuctif
21	7	1.9	311	MPCP_YEAST	P23641 saccharomyc
22	7	1.9	423	ENO_METJA	O60173 methanococc
23	7	1.9	445	XVLA_BACSU	P04788 bacillus su
24	7	1.9	516	GUX1_PCHAC	P13660 phanerochee
25	7	1.9	520	CMCH_NOCCLA	Q51080 noceardia la
26	7	1.9	532	GSI_NEUCR	P38678 neurospora
27	7	1.9	677	Y593_TREPA	O83602 treponema p
28	7	1.9	860	CHI2_COCCIM	P54197 coccidioid
29	7	1.9	904	Y002_CAERL	O09228 caenorhabdi
30	7	1.9	1018	STJ1_MOUSE	O99041 mus musculu
31	7	1.9	1019	STJ1_HUMAN	O99041 mus sapien
32	7	1.9	1026	BGAL_STRTR	P23989 streptococc
33	7	1.9	1043	SYL_METTH	O27428 methanobact

34	7	1.9	1044	1	SYL_METTH	P26499 methanobact
35	7	1.9	2485	1	PTND_HUMAN	O12923 homo sapien
36	6	1.6	25	1	PCN1_PACGO	P82423 pachycondyl
37	6	1.6	25	1	PCW2_PACGO	P82424 pachycondyl
38	6	1.6	71	1	YAAA_BACSU	P05650 bacillus su
39	6	1.6	76	1	B811_SCHCO	P78742 schizophyll
40	6	1.6	79	1	PSK3_ARATH	O9m200 arabidopsis
41	6	1.6	83	1	PRRP_RAT	P81378 rattus norv
42	6	1.6	83	1	RL23_HALLA	O06842 halobacteri
43	6	1.6	87	1	YIMC_BPPH1	P10435 bacterioph
44	6	1.6	88	1	PMRD_ECOLI	P37590 escherichia
45	6	1.6	88	1	Y40L_RHISN	P55597 rhizobium s

#### ALIGNMENTS

RESULT	1	STANDARD	PRT	353 AA.
AC	Q9N022; Q9B0G4;			
DT	01-MAR-2002 (Rel. 41, Created)			
DT	01-MAR-2002 (Rel. 41, Last sequence update)			
DT	01-MAR-2002 (Rel. 41, Last annotation update)			
DE	1-acyl-sn-glycerol-3-phosphate acyltransferase epsilon (EC 2.3.1.51)			
DE	(1-AGP acyltransferase 5) (1-AGPAT 5) (Lysophosphatidic acid			
DE	O-acyltransferase 5).			
GN	AGPAT5.			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.			
OX	NCBI_TaxID=9606;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RT	Leung D.W.;			
RT	"Cloning and expression of LPAAT-epsilon.";			
RT	Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.			
RN	[2]			
RP	SEQUENCE FROM N.A.			
RA	Cohen D., Chumakov I., Blumenfeld M., Bougueleret L.;			
RA	Patent number WO9932644, 01-JUL-1999.			
RN	[3]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=Amnydata;			
RX	MEDLINE=21154917; PubMed=11230166;			
RA	Wiedmann S., Well B., Wellenreuther R., Gassenhuber J., Glassl S.,			
RA	Ansoerge W., Boecker M., Bloeker H., Bauersachs S., Blum H.,			
RA	Lauber J., Duesterhoeft A., Beyer A., Koehrer K., Strack N.,			
RA	Mewes H.-W., Othenmaier B., Obermaier B., Tampe J., Heubner D.,			
RA	Wambolt R., Korn B., Klein M., Poustka A.;			
RT	"Towards a catalog of human genes and proteins: sequencing and			
RT	analysis of 500 novel complete protein coding human cDNAs.";			
RT	Genome Res. 11:422-435(2001).			
RN	[4]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=Placenta;			
RA	Isogai T., Ota T., Hayashi K., Sugiyama T., Otsuki T., Suzuki Y.,			
RA	Nishikawa T., Nagai K., Sugano S., Shiraori A., Sudo H.,			
RA	Tagakatsuna M., Hosoliri T., Kaku Y., Kodaira H., Kondo H., Sugawara M.,			
RA	Wakatsashi M., Chiba Y., Ishida S., Murakawa K., Ono Y., Takiguchi S.,			
RA	Watanabe S., Kimura K., Murakami K., Ishii S., Kawai Y., Saito K.,			
RA	Yamamoto J., Makatsus A., Nakamura Y., Nagahari K., Masuno Y.,			
RA	Nimono K., Iwayanagi T.;			
RT	"NEBD human cDNA sequencing project.";			
RT	Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.			
CC	-1- FUNCTION: CONVERTS LYOPHOSPHATIDIC ACID (LPA) INTO PHOSPHATIDIC			
CC	ACID BY INCORPORATING ACYL MOIETY AT THE 2 POSITION (BY			
CC	SIMILARITY).			
CC	-1- CATALYTIC ACTIVITY: Acyl-CoA + 1-acyl-sn-glycerol 3-phosphate -			
CC	CoA + 1,2-diacyl-sn-glycerol 3-phosphate.			
CC	-1- PATHWAY: SECOND STEP IN DE NOVO PHOSPHOLIPID BIOSYNTHESIS.			
CC	-1- SUBCELLULAR LOCATION: Integral membrane protein (potential).			

CC -1- SIMILARITY: BELONGS TO THE 1-ACYL-SN-GLYCEROL-3-PHOSPHATE  
 CC ACYLTRANSFERASE FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL: AF375789; AAK54809.1; ALT\_INIT.  
 DR EMBL: AL136587; CAB65522.1; ALT\_INIT.  
 DR EMBL: AK002072; BAB92009.1; -  
 DR InterPro: IPR002123; Acyltransferase.  
 DR Pfam: PF01553; Acyltransferase; 1.  
 KW Phospholipid biosynthesis; Transferase; Acyltransferase;  
 KW Transmembrane.  
 FT TRANSMEM. 7 23 POTENTIAL.  
 FT TRANSMEM. 44 66 POTENTIAL.  
 FT TRANSMEM. 334 351 POTENTIAL.  
 FT CONFLICT 145 145 L -> V (IN REF. 2).  
 SQ SEQUENCE 353 AA; 40813 MW; A05B1FA246CE1B64 CRC64;

Query Match 97.0%; Score 353; DB 1; Length 353;  
 Best Local Similarity 100.0%; Pred. No. 0;  
 Matches 353; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 12 MRYLLPSVLLGTAPTYLVANGVRLLSAFLPARFYQALDRLKCYVOSMLFEFENTG 71  
 DB 1 MRYLLPSVLLGTAPTYLVANGVRLLSAFLPARFYQALDRLKCYVOSMLFEFENTG 60  
 QY 72 VOILLYGDLPRKNENIYYLANHSTVDIVADILAIRONALGHRYVKEGLKMLPLYGC 131  
 DB 61 VOILLYGDLPRKNENIYYLANHSTVDIVADILAIRONALGHRYVKEGLKMLPLYGC 120  
 QY 132 YFAOHGIGYVRSKAFENKEMRNKLOSYVDAGTPMYLVIFPEGTRVNEQTKVLASQAF 191  
 DB 121 YFAOHGIGYVRSKAFENKEMRNKLOSYVDAGTPMYLVIFPEGTRVNEQTKVLASQAF 180  
 QY 192 AAOGLGLAVLKHVLTPIRIKATHVAFDCKKNYDAITYDVYVEGDDGQRESPTMTFFL 251  
 DB 181 AAOGLGLAVLKHVLTPIRIKATHVAFDCKKNYDAITYDVYVEGDDGQRESPTMTFFL 240  
 QY 252 CKECPKIHIDRIIDKRVPEEQEHMRRLHEFEIKDKMLIEFYESPDPERRRRPFCKS 311  
 DB 241 CKECPKIHIDRIIDKRVPEEQEHMRRLHEFEIKDKMLIEFYESPDPERRRRPFCKS 300  
 QY 312 VNSKLSIKTKLPSMLILSGLTAGMLMTDAGRKLVTYNTWYIGTLGCLMTVTKA 364  
 DB 301 VNSKLSIKTKLPSMLILSGLTAGMLMTDAGRKLVTYNTWYIGTLGCLMTVTKA 353

RESULT 2  
 PLSC\_MOUSE STANDARD; PRT; 354 AA.  
 AC 09D188;  
 DT 01-MAR-2002 (Rel. 41, Created)  
 DT 01-MAR-2002 (Rel. 41, Last sequence update)  
 DT 01-MAR-2002 (Rel. 41, Last annotation update)  
 DE 1-acyl-sn-glycerol-3-phosphate acyltransferase epsilon (EC 2.3.1.51)  
 DE (1-AGP acyltransferase 5) (1-AGPAT 5) (lysophosphatidic acid  
 DE acyltransferase-epsilon) (LPAAT-epsilon) (1-acylglycerol-3-phosphate  
 DE O-acyltransferase 5).  
 GN AGPAT5.  
 OS Mus musculus (mouse).  
 OC Eukaryota; Metazoa; Chordata; Cranialata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=C57BL/6J; TISSUE=Embryo;

EX MEDLINE=21085660; PubMed=11217851;  
 RA Kawai J., Shinaigawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,  
 RA Aikawa T., Hara A., Fukunishi Y., Kono H., Adachi J., Fukuda S.,  
 RA Akawa K., Izawa K., Nishi K., Kiyosawa H., Kondo S., Yamana K.,  
 RA Saito T., Okazaki Y., Gojobori T., Bono H., Katsukawa T., Saito R.,  
 RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,  
 RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,  
 RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,  
 RA Schirral L.M., Staehli F., Suzuki R., Tomita M., Wagner L., Washio T.,  
 RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,  
 RA Blake J., Boffelli D., Bojunga N., Carinci P., de Bonaldo M.F.,  
 RA Brownstein M.J., Bull C., Fletcher C., Fujita M., Gariboldi M.,  
 RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,  
 RA Lyons P., Marchionni L., Mashima J., Mazzerelli J., Mombaerts P.,  
 RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,  
 RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,  
 RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whitaker C., Wilmink L.,  
 RA Wyshaw-Boris A., Yoshida K., Hasegawa Y., Kawai J., Kohzuki S.,  
 RA Hayashizaki Y.,  
 RA \*Functional annotation of a full-length mouse cDNA collection.\*;  
 RT Nature 409:685-690(2001).  
 CC -1- FUNCTION: CONVERTS LYSOPHOSPHATIDIC ACID (LPA) INTO PHOSPHATIDIC  
 CC ACID BY INCORPORATING ACYL MOIETY AT THE 2 POSITION (BY  
 CC SIMILARITY).  
 CC -1- CATALYTIC ACTIVITY: Acyl-CoA + 1-acyl-sn-glycerol 3-phosphate =  
 CC CoA + 1,2-diacyl-sn-glycerol 3-phosphate.  
 CC -1- PATHWAY: SECOND STEP IN DE NOVO PHOSPHOLIPID BIOSYNTHESIS.  
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein (Potential).  
 CC -1- SIMILARITY: BELONGS TO THE 1-ACYL-SN-GLYCEROL-3-PHOSPHATE  
 CC ACYLTRANSFERASE FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL: AK003649; BAB22915.1; -  
 DR MGD: MGI:1915880; Agpat5.  
 DR InterPro: IPR002123; Acyltransferase.  
 DR Pfam: PF01553; Acyltransferase; 1.  
 KW Phospholipid biosynthesis; Transferase; Acyltransferase;  
 KW Transmembrane.  
 FT TRANSMEM. 7 29 POTENTIAL.  
 FT TRANSMEM. 44 66 POTENTIAL.  
 FT TRANSMEM. 335 352 POTENTIAL.  
 SQ SEQUENCE 354 AA; 40943 MW; 447EEF924B91E800 CRC64;

Query Match 8.8%; Score 32; DB 1; Length 354;  
 Best Local Similarity 100.0%; Pred. No. 16-24;  
 Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 185 LSASQAFPAORGGLAVLKHVLTPIRIKATHVAFD 216  
 DB 174 LSASQAFPAORGGLAVLKHVLTPIRIKATHVAFD 205  
 RESULT 3  
 PLSC\_HELPU STANDARD; PRT; 237 AA.  
 AC 09ZJNB;  
 DT 16-OCT-2001 (Rel. 40, Created)  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE 1-acyl-sn-glycerol-3-phosphate acyltransferase (EC 2.3.1.51) (1-AGP  
 DE acyltransferase) (1-AGPAT) (lysophosphatidic acid acyltransferase)  
 DE (LPAAT).  
 GN PLSC OR JHP1267.  
 OS Helicobacter pylori J99 (Campylobacter pylori J99).  
 OC Bacteria; Proteobacteria; epsilon subdivision; Helicobacter group;

```

OC Helicobacter.
OX NCBI_TaxID=85963;
RN [1]
RP SEQUENCE FROM N.A.
RA MEDLINE=99120557; PubMed=9923682;
RA Alm R.A., Ling L.-S.L., Mofr D.T., King B.L., Brown E.D., Doig P.C.,
RA Smith D.R., Noonan B., Guild B.C., deJonge B.L., Camel G.,
RA Tummulo P.J., Caruso A., Uria-Nickelsen M., Mills D.M., Ives C.,
RA Gibson R., Merberg D., Mills S.D., Jiang Q., Taylor D.E., Vovis G.F.,
RA Trust T.J.;
RT "Genomic sequence comparison of two unrelated isolates of the human
RT gastric pathogen Helicobacter pylori."
RL Nature 397:176-180(1999).
CC -1- FUNCTION: CONVERTS LYSOPHOSPHATIDIC ACID (LPA) INTO PHOSPHATIDIC
CC ACID BY INCORPORATING ACYL MOIETY AT THE 2 POSITION.
CC -1- CATALYTIC ACTIVITY: Acyl-CoA + 1-acyl-sn-glycerol 3-phosphate =
CC COA + 1,2-diacyl-sn-glycerol 3-phosphate.
CC -1- PATHWAY: SECOND STEP IN DE NOVO PHOSPHOLIPID BIOSYNTHESIS.
CC -1- SUBCELLULAR LOCATION: Inner membrane-associated (Potential).
CC -1- SIMILARITY: BELONGS TO THE 1-ACYL-SN-GLYCEROL-3-PHOSPHATE
CC ACYLTRANSFERASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sid.ch).
CC -----
CC EMBL: AE001550; AAD06852.1; -
CC DR InterPro: IPR002123; Acyltransferase.
CC DR Pfam: PF01553; Acyltransferase: 1.
CC KW Phospholipid biosynthesis; Transferase; Acyltransferase;
CC Inner membrane; Complete proteome.
CC SEQUENCE 237 AA; 27188 MW; E10F517D42A1731F CRC64;

Query Match 2.5%; Score 9; DB 1; Length 237;
Best Local Similarity 100.0%; Pred. No. 0.22;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 168 LVIFPECTR 176
ID |||||||
DB 142 LVIFPECTR 150

RESULT 4
PLSC_HELPY STANDARD; PRT; 240 AA.
AC 025903:
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE 1-acyl-sn-glycerol-3-phosphate acyltransferase (EC 2.3.1.51) (1-AGP
DE acyltransferase) (1-AGPAT) (Lysophosphatidic acid acyltransferase)
DE (LIPART).
GN PLSC OR HP1348.
OS Helicobacter pylori (Campylobacter pylori).
OC Bacteria; Proteobacteria; epsilon subdivision; Helicobacter group;
OC Helicobacter.
OX NCBI_TaxID=210;
RN [1]
RP SEQUENCE FROM N.A.
RA STRAIN=26695 / ATCC 700392;
RA MEDLINE=97394467; PubMed=9252185;
RA Tomb J.F., White O., Kerlavage A.R., Clayton R.A., Sutton G.G.,
RA Fleischmann R.D., Ketchum K.A., Klenk H.-P., Gill S., Dougherty B.A.,
RA Nelson K., Quackenbush J., Zhou L., Kirkness E.F., Peterson S.,
RA Loftus B., Richardson D., Dodson R., Khalak H.G., Glodek A.,
RA McKenney K., Fitzgerald L.M., Lee N., Adams M.D., Hickey E.K.,
RA Berg D.E., Gocayne J.D., Uterback T.R., Peterson J.D., Kelley J.M.,
RA Cotton M.D., Weidman J.M., Fujii C., Bowman C., Watthey L., Wallin E.,

```

```

RA Hayes W.S., Borodovsky M., Karp P.D., Smith H.O., Fraser C.M.,
RA Venter J.C.;
RT "The complete genome sequence of the gastric pathogen Helicobacter
RT pylori."
RL Nature 388:539-547(1997).
CC -1- FUNCTION: CONVERTS LYSOPHOSPHATIDIC ACID (LPA) INTO PHOSPHATIDIC
CC ACID BY INCORPORATING ACYL MOIETY AT THE 2 POSITION.
CC -1- CATALYTIC ACTIVITY: Acyl-CoA + 1-acyl-sn-glycerol 3-phosphate =
CC COA + 1,2-diacyl-sn-glycerol 3-phosphate.
CC -1- PATHWAY: SECOND STEP IN DE NOVO PHOSPHOLIPID BIOSYNTHESIS.
CC -1- SUBCELLULAR LOCATION: Inner membrane-associated (Potential).
CC -1- SIMILARITY: BELONGS TO THE 1-ACYL-SN-GLYCEROL-3-PHOSPHATE
CC ACYLTRANSFERASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: AE000636; AAD08393.1; -
CC DR TIGR: HP1348; -
CC DR InterPro: IPR002123; Acyltransferase.
CC DR Pfam: PF01553; Acyltransferase: 1.
CC KW Phospholipid biosynthesis; Transferase; Acyltransferase;
CC Inner membrane; Complete proteome.
CC SEQUENCE 240 AA; 27745 MW; 22BD5D0EB190BBD CRC64;

Query Match 2.5%; Score 9; DB 1; Length 240;
Best Local Similarity 100.0%; Pred. No. 0.22;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 168 LVIFPECTR 176
ID |||||||
DB 142 LVIFPECTR 150

RESULT 5
V187_BP73 STANDARD; PRT; 83 AA.
AC P10302:
DT 01-MAR-1989 (Rel. 10, Created)
DT 01-MAR-1989 (Rel. 10, Last sequence update)
DT 01-MAR-1989 (Rel. 10, Last annotation update)
DE Gene 18.7 protein.
GN 18.7.
OS Bacteriophage T3.
OC Viruses; dsDNA viruses, no RNA stage; Caudovirales; Podoviridae;
OC T7-like phages.
OX NCBI_TaxID=10759;
RN [1]
RP SEQUENCE FROM N.A.
RA MEDLINE=86209997; PubMed=3010556;
RA Yamada M., Fujisawa H., Kato H., Hamada K., Mingawa T.;
RT "Cloning and sequencing of the genetic right end of bacteriophage T3
RL virology 151:350-361(1986).
RN [2]
RP ERRATUM.
RA Yamada M., Fujisawa H., Kato H., Hamada K., Mingawa T.;
RL Virology 154:246-246(1986).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----

```

DR EMBL: M14784; AAA92527.1; -  
 DR PIR: E23476; WBBPT3.  
 SQ SEQUENCE 83 AA; 9393 MW; 63128984CBAF0531 CRC64;

Query Match  
 Best Local Similarity 100.0%; Score 8; DB 1; Length 83;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 316 LSIKKTLP 323  
 |||||||  
 Db 22 LSIKKTLP 29

RESULT 6  
 V187\_BP77 STANDARD; PRT; 83 AA.  
 AC P03786;  
 DT 21-JUL-1986 (Rel. 01, Created)  
 DT 21-JUL-1986 (Rel. 01, Last sequence update)  
 DT 01-MAR-1989 (Rel. 10, Last annotation update)  
 DE Gene 18.7 protein.  
 GN 18.7.  
 OS Bacteriophage T7.  
 OC Viruses; dsDNA viruses, no RNA stage; Caudovirales; Podoviridae;  
 OC T7-like phages.  
 OX NCBI\_TaxID=10760;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=83241725; PubMed=6864790;  
 RA Dunn J.J., Studier F.W.;  
 RT "Complete nucleotide sequence of bacteriophage T7 DNA and the  
 RT locations of T7 genetic elements.";  
 RL J. Mol. Biol. 166:477-535(1983).  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL: V01146; CAA24439.1; -  
 DR PIR: A04413; WBBPG7.  
 DR PIR: S42337; S42337.  
 SQ SEQUENCE 83 AA; 9326 MW; B3DB055DE79A53C5 CRC64;

Query Match  
 Best Local Similarity 100.0%; Score 8; DB 1; Length 83;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 316 LSIKKTLP 323  
 |||||||  
 Db 22 LSIKKTLP 29

RESULT 7  
 YD18\_YEAST STANDARD; PRT; 396 AA.  
 AC Q12185;  
 DT 15-JUL-1999 (Rel. 38, Created)  
 DT 15-JUL-1999 (Rel. 38, Last sequence update)  
 DT 15-JUL-1999 (Rel. 38, Last annotation update)  
 DE Hypothetical 45.9 kDa protein in KCS1-GCV1 intergenic region.  
 GN YOR018C OR YD9335.04C OR PZE596.  
 OS Saccharomyces cerevisiae (Baker's yeast).  
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;  
 OC Saccharomycetales; Saccharomycetaceae; Saccharomyces.  
 OX NCBI\_TaxID=49332;  
 RN [1]  
 RP SEQUENCE FROM N.A.

RC STRAIN=S288C / AB972;  
 RA Dedman K., Brown D., Hamlyn N., Bowman S., Barrell B.G.,  
 RA Rajandream M.A.;  
 RL Submitted (MAY-1995) to the EMBL/GenBank/DBJ databases.

Query Match  
 Best Local Similarity 100.0%; Score 8; DB 1; Length 396;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 222 LDAIVDVT 229  
 |||||||  
 Db 265 LDAIVDVT 272

RESULT 8  
 HSCA\_BUCAL STANDARD; PRT; 611 AA.  
 AC P57660;  
 DT 16-OCT-2001 (Rel. 40, Created)  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Chaperone protein hscA homolog.  
 GN HSCA OR BU605.  
 OS Buchnera aphidicola (subsp. Acyrthosiphon pisum) (Acyrthosiphon pisum  
 OS symbiotic bacterium).  
 OC Bacteria; Proteobacteria; gamma subdivision; Buchnera.  
 OX NCBI\_TaxID=118099;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=TOKYO 1998;  
 RX MEDLINE=20445173; PubMed=10993077;  
 RA Shigenobu S., Matanabe H., Hattori M., Sakaki Y., Ishikawa H.;  
 RT "Genome sequence of the endocellular bacterial symbiont of aphids  
 RT Buchnera sp. APS.";  
 RL Nature 407:81-86(2000).  
 CC -1- FUNCTION: PROBABLE CHAPERONE. HAS A LOW INTRINSIC ATPASE ACTIVITY  
 CC WHICH IS MARKEDLY STIMULATED BY HSCB (BY SIMILARITY).  
 CC -1- SIMILARITY: BELONGS TO THE HEAT SHOCK PROTEIN 70 FAMILY.  
 CC -----



CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL: AP001119; BAB13289.1; -  
 DR Interpro: IPR01023; HSP70.  
 DR Pfam: PF00012; HSP70.1.  
 DR PRINTS: PR00301; HEATSHOCK70.  
 DR PROSITE: PS00297; HSP70\_1; 1.  
 DR PROSITE: PS00329; HSP70\_2; 1.  
 DR PROSITE: PS01036; HSP70\_3; 1.  
 KM Chaperone; ATP-binding; Complete proteome.  
 SQ SEQUENCE 611 AA; 69837 MW; 32B567C53073082A CRC64;

Query Match 2.2%; Score 8; DB 1; Length 611;  
 Best Local Similarity 100.0%; Pred. No. 5.1;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 13 RYLPSV 20  
 |||||  
 DB 42 RYLPSV 49

## RESULT 9

ID YJDO\_ECOLI STANDARD; PRT; 57 AA.  
 AC P58038;  
 DT 16-OCT-2001 (Rel. 40, Last created)  
 DT 16-OCT-2001 (Rel. 41, Last annotation update)  
 DT 01-MAR-2002 (Rel. 41, Last annotation update)  
 DE Hypothetical protein yj40.  
 GN YJDO OR B4128.1 OR Z5731 OR EC55110.1.  
 OS Escherichia coli, and  
 OS Escherichia coli O157:H7.  
 OC Escherichia; Proteobacteria; gamma subdivision; Enterobacteriaceae;  
 OC Escherichia.  
 OX NCBI\_Taxid=562, 83334;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC MEDLINE=97426617; PubMed=9278503;  
 RA Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,  
 RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,  
 RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,  
 RA Mau B., Shao Y.;  
 RT "The complete genome sequence of Escherichia coli K-12.";  
 RL Science 277:1453-1474(1997).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=O157:H7 / EDL933 / ATCC 700927;  
 RX MEDLINE=21074935; PubMed=11206551;  
 RA Perna N.T., Plunkett G. III, Burland V., Mau B., Glasner J.D.,  
 RA Rose D.J., Mayhew G.F., Evans P.S., Gregor J., Kirkpatrick H.A.,  
 RA Postai G., Hackett J., Klink S., Boutin A., Shao Y., Miller L.,  
 RA Grodock E.J., Davis N.W., Lim A., Dimalanta E.T., Potamotis K.,  
 RA Apodaca J., Anantharaman T.S., Lin J., Yen G., Schwartz D.C.,  
 RA Welch R.A., Blattner F.R.;  
 RT "Genome sequence of enterohaemorrhagic Escherichia coli O157:H7.";  
 RL Nature 409:529-533(2001).  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=O157:H7 / RIMD 0509952;  
 RX MEDLINE=21156231; PubMed=11258796;  
 RA Hayashi T., Makino K., Ohnishi M., Kurokawa K., Ishii K., Yokoyama K.,  
 RA Han C.-G., Ohtsubo E., Nakayama K., Murata T., Tanaka M., Tode T.,  
 RA Iida T., Takami H., Honda T., Sasaki C., Ogasawara N., Yasunaga T.,  
 RA Kuhara S., Shiba T., Hattori M., Shinagawa H.;  
 RT "Complete genome sequence of enterohemorrhagic Escherichia coli

RT O157:H7 and genomic comparison with a laboratory strain K-12.";  
 RL DNA Res. 8:11-22(2001).  
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein (potential).  
 CC -1- SIMILARITY: STRONG, TO E.COLI YDCX.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----

DR EMBL: AE000485; -; NOT\_ANNOTATED\_CDS.  
 DR EMBL: AE005646; AAG59328.1; -;  
 DR EcoGene: EGI4342; yj40.  
 DR Hypothetical protein; Transmembrane; Complete proteome.  
 KW Hypothetical protein; Transmembrane; Complete proteome.  
 FT TRANSMEM 37 57  
 FT POTENTIAL.  
 SQ SEQUENCE 57 AA; 6555 MW; A3670A19500F75D6 CRC64;

Query Match 1.9%; Score 7; DB 1; Length 57;  
 Best Local Similarity 100.0%; Pred. No. 7;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 36 RLISAF 42  
 |||||  
 DB 33 RLISAF 39

## RESULT 10

ID PUR1\_LACCA STANDARD; PRT; 194 AA.  
 AC P35853;  
 DT 01-JUN-1994 (Rel. 29, Created)  
 DT 01-FEB-1996 (Rel. 33, Last annotation update)  
 DT 30-MAY-2000 (Rel. 39, Last annotation update)  
 DE Amidophosphoribosyltransferase precursor (EC 2.4.2.14) (glutamine  
 DE phosphoribosylpyrophosphate amidotransferase) (ATase) (GPATase)  
 DE (Fragment).  
 GN PURP.  
 OS Lactobacillus casei.  
 OC Bacteria; Firmicutes; Bacillus/Clostridium group; Lactobacillaceae;  
 OC Lactobacillus.  
 OX NCBI\_Taxid=1582;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=93012962; PubMed=1398079;  
 RA Gu Z.-M., Martindale D.W., Lee B.H.;  
 RT "Isolation and complete sequence of the purL gene encoding PCAM  
 RT synthase II in Lactobacillus casei.";  
 RL Gene 119:123-126(1992).  
 RN [2]  
 RP ERRATUM.  
 RX MEDLINE=94040790; PubMed=8224889;  
 RA Gu Z.-M., Martindale D.W., Lee B.H.;  
 RL Gene 133:147-147(1993).  
 CC -1- CATALYTIC ACTIVITY: 5-phospho-beta-D-ribose-5-phosphate +  
 CC L-glutamate -> L-glutamine + 5-phospho-alpha-D-ribose 1-diphosphate  
 CC + H(2)O.  
 CC -1- PATHWAY: FIRST STEP IN DE NOVO PURINE BIOSYNTHESIS  
 CC -1- SIMILARITY: THE GATASE DOMAIN BELONGS TO TYPE-2 GLUTAMINE  
 CC AMIDOTRANSFERASES.  
 CC -1- SIMILARITY: IN THE C-TERMINAL SECTION, BELONGS TO THE  
 CC PURINE/PYRIMIDINE PHOSPHORIBOSYLTRANSFERASE FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 -----  
 DR EMBL: M85265; AAC36948.1; -  
 DR PIR: PC1136; PC1136.  
 DR HSSP: P00497; 1A00.  
 DR MEROPS: C44.001; -  
 DR InterPro: IPR000583; GATase\_2.  
 DR InterPro: IPR002375; Pur\_Pyr\_Pr\_transf.  
 DR Pfam: PF00310; GATase\_2; 1.  
 DR PROSITE: PS00103; PUR\_PYR\_PR\_TRANSFERR; PARTIAL.  
 DR PROSITE: PS00443; GATASE\_TYPE\_II; 1.  
 DR Purine biosynthesis; Transferrase; Glycosyltransferase.  
 KW PROPEP 1 1..  
 FT CHAIN 12 >194  
 FT ACT\_SITE 12 12  
 FT NON\_TER 194 194  
 FT SEQUENCE 194 AA; 21144 MW; 4A788BC5365D5EC CRC64;

Query Match 1.9%; Score 7; DB 1; Length 194;  
 Best Local Similarity 100.0%; Pred. No. 20;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 111 ALGHVRY 117  
 |||||  
 Db 79 ALGHVRY 85

RESULT 11  
 HBGF\_CERAE STANDARD; PRT; 208 AA.  
 ID HBGF\_CERAE 009118;  
 DT 01-FEB-1996 (Rel. 33, Created)  
 DT 01-FEB-1996 (Rel. 33, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Heparin-binding EGF-like growth factor precursor (HB-EGF) (HBEGF)  
 DE (Diphtheria toxin receptor) (DT-R).  
 DE DTR OR HEGFL.  
 OS Cercopithecus aethiops (Green monkey) (Grivet).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;  
 OC Cercopithecinae; Cercopithecus.  
 NC NCBL\_TaxID=9534;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=92298386; Pubmed=1606612;  
 RA Neglich J.G., Metherall J.E., Russel D.W., Eldels L.;  
 RT "Expression cloning of a diphtheria toxin receptor: Identity with a  
 heparin-binding EGF-like growth factor precursor.";  
 RL Cell 69:1051-1061(1992).  
 RN [2]  
 RP TOXIN-BINDING DOMAIN.  
 RX MEDLINE=95126975; Pubmed=7826391;  
 RA Hooper K.P., Eldels L.;  
 RT "Localization of a critical diphtheria toxin-binding domain to the C-  
 terminus of the mature heparin-binding EGF-like growth factor region  
 of the diphtheria toxin receptor.";  
 RL Biochem. Biophys. Res. Commun. 206:710-717(1995).  
 CC -1- FUNCTION: MAY BE INVOLVED IN MACROPHAGE-MEDIATED CELLULAR  
 PROLIFERATION. IT IS MITOGENIC FOR FIBROBLASTS AND SMOOTH MUSCLE  
 BUT NOT ENDOTHELIAL CELLS. IT IS ABLE TO BIND EGF RECEPTORS WITH  
 HIGHER AFFINITY THAN EGF ITSELF AND IS A FAR MORE POTENT MITOGEN  
 FOR SMOOTH MUSCLE CELLS THAN EGF (BY SIMILARITY).  
 CC -1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN. MATURE HB-EGF IS  
 RELEASED INTO THE EXTRACELLULAR SPACE AND PROBABLY BINDS TO A  
 RECEPTOR (BY SIMILARITY).  
 CC -1- PTM: O-GLYCOSYLATED (BY SIMILARITY).  
 CC -1- SIMILARITY: CONTAINS 1 EGF-LIKE DOMAIN.  
 -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 -----

DR EMBL: M93012; -; NOT\_ANNOTATED\_CDS.  
 DR HSSP: Q99075; 1XDT.  
 DR InterPro: IPR000561; EGF-like.  
 DR Pfam: PF00008; EGF; 1.  
 DR SMART: SM00181; EGF; 1.  
 DR PROSITE: PS00022; EGF\_1; 1.  
 DR PROSITE: PS01186; EGF\_2; 1.  
 DR Growth factor; Heparin-binding; Signal; EGF-like domain;  
 KW Transmembrane; Glycoprotein; Receptor.  
 FT SIGNAL 1 19  
 FT PROPEP 20 62  
 FT CHAIN 63 148  
 FT PROPEP 149 208  
 FT DOMAIN 20 160  
 FT TRANSMEM 161 184  
 FT DOMAIN 185 208  
 FT DOMAIN 104 144  
 FT CARBOHYD 75 75  
 FT CARBOHYD 85 85  
 FT DISULFID 108 121  
 FT DISULFID 116 132  
 FT DISULFID 134 143  
 FT SEQUENCE 208 AA; 22985 MW; 8D108289A0485AE9 CRC64;

Query Match 1.9%; Score 7; DB 1; Length 208;  
 Best Local Similarity 100.0%; Pred. No. 21;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 15 LTPSVVL 21  
 |||||  
 Db 3 LTPSVVL 9

RESULT 12  
 HBGF\_HUMAN STANDARD; PRT; 208 AA.  
 ID HBGF\_HUMAN 099075;  
 DT 01-JUN-1994 (Rel. 29, Created)  
 DT 01-JUN-1994 (Rel. 29, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Heparin-binding EGF-like growth factor precursor (HB-EGF) (HBEGF)  
 DE (Diphtheria toxin receptor) (DT-R).  
 DE DTR OR HEGFL.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
 NC NCBL\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A. AND SEQUENCE OF 73-93.  
 RX TISSUE-Macrophage;  
 RX MEDLINE=91157008; Pubmed=1840698;  
 RA Hgashiyama S., Abraham J.A., Miller J., Fiddes J.C., Klagsbrun M.;  
 RT "A heparin-binding growth factor secreted by macrophage-like cells  
 that is related to EGF.";  
 RL Science 251:936-939(1991).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RA Kimmmerly W., Bondoc M., Cheng J., Connolly K.S., Gunning K.M.,  
 RA Kadner K., Miguel T., Miller C., Piltuck S., Pollard M., Rojeski H.,  
 RA Subramanian S., Martin C.H.;  
 RL Submitted (Apr-1998) to the EMBL/GenBank/DBJ databases.  
 RN [3]  
 RP SEQUENCE OF 63-141 AND 143-148, AND CARBOHYDRATE-LINKAGE SITES.  
 RX TISSUE-Histiocytic lymphoma;  
 RX MEDLINE=92210396; Pubmed=1556128;  
 RA Hgashiyama S., Lau K., Besner G.E., Abraham J.A., Klagsbrun M.;

RT "Structure of heparin-binding EGF-like growth factor. Multiple forms, primary structure, and glycosylation of the mature protein.";  
 RL J. Biol. Chem. 267:6205-6212(1992).  
 RN [4]  
 RP TOXIN-BINDING DOMAIN  
 RX MEDLINE-95138082; PubMed-7836353;  
 RA Mitamura T., Higashiyama S., Taniguchi N., Klagsbrun M., Mekada E.;  
 RT "Diphtheria toxin binds to the epidermal growth factor (EGF)-like domain of human heparin-binding EGF-like growth factor/diphtheria toxin receptor and inhibits specifically its mitogenic activity.";  
 RL J. Biol. Chem. 270:1015-1019(1995).  
 RN [5]  
 RP X-RAY CRYSTALLOGRAPHY (2.65 ANGSTROMS) OF 73-147 IN COMPLEX WITH TOX.  
 RX MEDLINE-98324089; PubMed-9659904;  
 RA Louie G.V., Yang W., Bowman M.E., Choe S.;  
 RT "Crystal structure of the complex of diphtheria toxin with an extracellular fragment of its receptor.";  
 RL Mol. Cell 1:67-78(1997).  
 CC -1- FUNCTION: MAY BE INVOLVED IN MACROPHAGE-MEDIATED CELLULAR PROLIFERATION. IT IS MITOGENIC FOR FIBROBLASTS AND SMOOTH MUSCLE BUT NOT ENDOTHELIAL CELLS. IT IS ABLE TO BIND EGF RECEPTORS WITH HIGHER AFFINITY THAN EGF ITSELF AND IS A FAR MORE POTENT MITOGEN FOR SMOOTH MUSCLE CELLS THAN EGF.  
 CC -1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN. MATURE HB-EGF IS RELEASED INTO THE EXTRACELLULAR SPACE AND PROBABLY BINDS TO A RECEPTOR.  
 CC -1- PTM: SEVERAL N-TERMINI HAVE BEEN IDENTIFIED BY DIRECT SEQUENCING. THE FORMS WITH N-TERMINI 63, 73 AND 74 HAVE BEEN TESTED AND FOUND TO BE BIOLOGICALLY ACTIVE.  
 CC -1- PTM: O-GLYCAN ATTACHMENT SITES WERE DETERMINED BY EDMAN DEGRADATION, O-GLYCANASE DIGEST SUGGESTS MUClN-TYPE GLYCOSYLATION (DONE IN HB-EGF PURIFIED FROM HISTIOCYTIC LYMPHOMA CELL LINE U-937).  
 CC -1- SIMILARITY: CONTAINS 1 EGF-LIKE DOMAIN.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sdb.ch/announce/> or send an email to [license@sdb-sib.ch](mailto:license@sdb-sib.ch)).  
 CC -----  
 CC EMBL: M60278; AAA35956.1; -  
 CC EMBL: AC004634; AAC15470.1; -  
 CC PIR: A38432; A38432  
 CC PDB: 1XDT; 25-FEB-98.  
 CC MIM: 126150; -  
 CC InterPro: IPR000561; EGF-like.  
 CC DR Pfam: PF00008; EGF; 1.  
 CC DR SMART: SM00181; EGF; 1.  
 CC DR PROSITE: PS00022; EGF\_1; 1.  
 CC DR PROSITE: PS01186; EGF\_2; 1.  
 CC Growth factor; Heparin-binding; Signal; EGF-like domain;  
 CC Transmembrane; Glycoprotein; Receptor; 3D-structure.  
 CC SIGNAL 1 19  
 CC PROPEP 20 62 OR 72, OR 73, OR 76, OR 81.  
 CC CHAIN 149 208 HEPARIN-BINDING EGF-LIKE GROWTH FACTOR.  
 CC PROPEP 149 208 C-TERMINAL (POTENTIAL).  
 CC DOMAIN 20 160 EXTRACELLULAR (POTENTIAL).  
 CC TRANSMEM 161 184 CYTOPLASMIC (POTENTIAL).  
 CC DOMAIN 185 208 POTENTIAL.  
 CC DOMAIN 104 144 EGF-LIKE.  
 CC DISULFID 108 121 BY SIMILARITY.  
 CC DISULFID 116 132 BY SIMILARITY.  
 CC CARBOHYD 134 143 BY SIMILARITY.  
 CC CARBOHYD 75 75 O-LINKED (GLNAC. . .).  
 CC CARBOHYD 85 85 O-LINKED (GLNAC. . .).  
 CC SEQUENCE 208 AA; 23067 MW; 2C43C9D1D8291B51 CRC64;

Query Match 1.9%; Score 7; DB 1; Length 208;  
 Best Local Similarity 100.0%; Pred. No. 21;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 15 LPPSVVL 21  
 |||||  
 Db 3 LPPSVVL 9  
 RESULT 13  
 HBGF\_PIG  
 ID HBGF\_PIG STANDARD; PRT; 208 AA.  
 AC 001580;  
 DT 01-FEB-1996 (Rel. 33, Created)  
 DT 30-MAY-2000 (Rel. 39, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Heparin-binding EGF-like growth factor precursor (HB-EGF) (HBEGF).  
 GN DTR OR HBGL.  
 OS Sus scrofa (Pig).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.  
 RN NCBI\_TaxID=9823;  
 RP SEQUENCE FROM N.A.  
 RA Pascall J.C.;  
 RL Submitted (NOV-1997) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP SEQUENCE OF 99-182 FROM N.A.  
 RC STRAIN=LARGE WHITE; TISSUE=Heart;  
 RX MEDLINE-93075016; PubMed-1445231;  
 RA Vaughan T.J., Pascall J.C., Brown K.D.;  
 RT "Tissue distribution of mRNA for heparin-binding epidermal growth factor";  
 RL Biochem. J. 287:681-684(1992).  
 CC -1- FUNCTION: MAY BE INVOLVED IN MACROPHAGE-MEDIATED CELLULAR PROLIFERATION. IT IS MITOGENIC FOR FIBROBLASTS AND SMOOTH MUSCLE BUT NOT ENDOTHELIAL CELLS. IT IS ABLE TO BIND EGF RECEPTORS WITH HIGHER AFFINITY THAN EGF ITSELF AND IS A FAR MORE POTENT MITOGEN FOR SMOOTH MUSCLE CELLS THAN EGF (BY SIMILARITY).  
 CC -1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN. MATURE HB-EGF IS RELEASED INTO THE EXTRACELLULAR SPACE AND PROBABLY BINDS TO A RECEPTOR (BY SIMILARITY).  
 CC -1- TISSUE SPECIFICITY: MACROPHAGE, MIDBRAIN, CEREBELLUM, HYPOTHALAMUS, CEREBRAL CORTEX, BUDBOURETHRAL GLAND, LUNG, HEART VENTRICLE, KIDNEY, SKIN, PROSTATE, SEMINAL VESICLE, TESTIS; AT LOW LEVELS IN LYMPH NODE, THYMUS, SPLEEN; NOT DETECTED IN PITUITARY, OLFACORY BULB, THYROID, DUODENUM, PANCREAS, LIVER, SUBMAXILLARY GLAND.  
 CC -1- PTM: O-GLYCOSYLATED (PROBABLE).  
 CC -1- SIMILARITY: CONTAINS 1 EGF-LIKE DOMAIN.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sdb.ch/announce/> or send an email to [license@sdb-sib.ch](mailto:license@sdb-sib.ch)).  
 CC -----  
 CC EMBL: Y15731; CAA75740.1; -  
 CC EMBL: X67295; CAA47709.1; -  
 CC HSSP: Q99075; 1XDT.  
 CC InterPro: IPR000561; EGF-like.  
 CC InterPro: IPR001336; EGF\_1.  
 CC Pfam: PF00008; EGF; 1.  
 CC PRINTS: PR00009; EGF\_TGF.  
 CC SMART: SM00181; EGF; 1.  
 CC PROSITE: PS00022; EGF\_1; 1.  
 CC PROSITE: PS01186; EGF\_2; 1.  
 CC Signal; Growth factor; Heparin-binding; EGF-like domain;  
 CC Transmembrane; Glycoprotein.  
 CC SIGNAL 1 23  
 CC PROPEP 24 62 POTENTIAL.  
 CC CHAIN 63 148 HEPARIN-BINDING EGF-LIKE GROWTH FACTOR.  
 CC PROPEP 149 208 C-TERMINAL (POTENTIAL).

FT DOMAIN 24 161 EXTRACELLULAR (POTENTIAL).  
 FT TRANSMEM 162 182 POTENTIAL.  
 FT DOMAIN 183 208 CYTOPLASMIC (POTENTIAL).  
 FT DOMAIN 104 144 EGF-LIKE.  
 FT DISULFID 108 121 BY SIMILARITY.  
 FT DISULFID 116 132 BY SIMILARITY.  
 FT DISULFID 134 143 BY SIMILARITY.  
 FT CARBOHYD 85 85 O-LINKED (GALNAC. . .) (BY SIMILARITY).  
 SQ SEQUENCE 208 AA: 22986 MW: 0A7DA97AE30C8967 CRC64;

Query Match 1.9%; Score 7; DB 1; Length 208;  
 Best Local Similarity 100.0%; Pred. No. 21;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 15 LPSVYL 21  
 |||||  
 DB 3 LPSVYL 9

RESULT 14  
 HBGF\_RAT  
 ID HBGF\_RAT STANDARD: PRT: 208 AA.  
 AC 006175;

DT 01-FEB-1995 (Rel. 31, Created)  
 DT 01-FEB-1995 (Rel. 31, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Heparin-binding EGF-like growth factor precursor (HB-EGF) (HBEGF).  
 DN DTR OR HBGF.

OS Rattus norvegicus (Rat).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
 ON NCBI\_TaxID=10116;

RP SEQUENCE FROM N.A.  
 RC STRAIN=SPRAGUE-DAWLEY; TISSUE=Macrophage;  
 RX MEDLINE=93135756; Pubmed=7678488;

RA Abraham J.A., Damm D., Hajdari A., Miller J., Klagsbrun M.,  
 RA Ezekowitz R.A.B.;

RT "Heparin-binding EGF-like growth factor: characterization of rat and  
 mouse cDNA clones, protein domain conservation across species, and  
 RT transcript expression in tissues."

RL Biochem. Biophys. Res. Commun. 190:125-133(1993).  
 CC -1- FUNCTION: MAY BE INVOLVED IN MACROPHAGE-MEDIATED CELLULAR  
 CC PROLIFERATION. IT IS MITOGENIC FOR FIBROBLASTS AND SMOOTH MUSCLE  
 CC BUT NOT ENDOTHELIAL CELLS. IT IS ABLE TO BIND EGF RECEPTORS WITH  
 CC HIGHER AFFINITY THAN EGF ITSELF AND IS A FAR MORE POTENT MITOGEN  
 CC FOR SMOOTH MUSCLE CELLS THAN EGF.

CC -1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN. MAJORE HB-EGF IS  
 CC RELEASED INTO THE EXTRACELLULAR SPACE AND PROBABLY BINDS TO A  
 CC RECEPTOR.  
 CC -1- TISSUE SPECIFICITY: MOST ABUNDANT IN SKELETAL MUSCLE, LUNG, SPLEEN  
 CC BRAIN AND HEART.

CC -1- PTM: O-GLYCOSYLATED (BY SIMILARITY).  
 CC -1- SIMILARITY: CONTAINS 1 EGF-LIKE DOMAIN.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC EMBL: L05489; AAA81780.1; .  
 CC PIR: JCI409; JCI409.

DR HSSP: Q99075; 1XDT.  
 DR InterPro: IPR000561; EGF-like.

DR Pfam: PF00008; EGF\_1.

DR SMART: SM00181; EGF\_1.

DR PROSITE: PS00022; EGF\_1; 1.

DR PROSITE: PS01186; EGF\_2; 1.

KW Signal; Growth factor; Heparin-binding; EGF-like domain;

KW Transmembrane; Glycoprotein.  
 FT SIGNAL 1 23 POTENTIAL.  
 FT PROPEP 24 62 BY SIMILARITY.  
 FT CHAIN 63 148 HEPARIN-BINDING EGF-LIKE GROWTH FACTOR.  
 FT PROPEP 149 208 C-TERMINAL (POTENTIAL).  
 FT DOMAIN 24 160 EXTRACELLULAR (POTENTIAL).  
 FT TRANSMEM 161 184 POTENTIAL.  
 FT DOMAIN 185 208 CYTOPLASMIC (POTENTIAL).  
 FT DISULFID 104 144 EGF-LIKE.  
 FT DISULFID 108 121 BY SIMILARITY.  
 FT DISULFID 116 132 BY SIMILARITY.  
 FT DISULFID 134 143 BY SIMILARITY.  
 FT CARBOHYD 85 85 O-LINKED (GALNAC. . .) (BY SIMILARITY).  
 SQ SEQUENCE 208 AA: 22843 MW: DDBD045E116D064C CRC64;

Query Match 1.9%; Score 7; DB 1; Length 208;  
 Best Local Similarity 100.0%; Pred. No. 21;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 15 LPSVYL 21  
 |||||  
 DB 3 LPSVYL 9

RESULT 15  
 HXB4\_MOUSE  
 ID HXB4\_MOUSE STANDARD: PRT: 250 AA.  
 AC P10284;

DT 01-MAR-1989 (Rel. 10, Created)  
 DT 01-MAR-1989 (Rel. 10, Last sequence update)  
 DT 01-OCT-1996 (Rel. 34, Last annotation update)  
 DE Homeobox protein Hox-B4 (Hox-2.6).  
 DN HOXB4 OR HOXB-4 OR HOX-2.6.

OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 ON NCBI\_TaxID=10090;

RP SEQUENCE FROM N.A.  
 RC MEDLINE=89091992; Pubmed=2463210;

RA Graham A., Papalopulu N., Lorimer J., McVey J.H., Tuddenham E.G.D.,  
 RA Krangel R.;

RT "Characterization of a murine homeo box gene, Hox-2.6, related to the  
 RT Drosophila Deformed gene."

CC -1- FUNCTION: SEQUENCE-SPECIFIC TRANSCRIPTION FACTOR WHICH IS PART OF  
 CC A DEVELOPMENTAL REGULATORY SYSTEM THAT PROVIDES CELLS WITH  
 CC SPECIFIC POSITIONAL IDENTITIES ON THE ANTERIOR-POSTERIOR AXIS.  
 CC -1- SUBCELLULAR LOCATION: Nuclear.  
 CC -1- SIMILARITY: BELONGS TO THE ANTP FAMILY OF HOMEBOX PROTEINS.  
 CC "DEFORMED" SUBFAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC EMBL: M36654; AAA37848.1; .

DR PIR: A31757; A31757.

DR HSSP: P02833; 9ANT.

DR TRANSFAC: T01728; .

DR MGD: MGT:96185; HOXB4

DR InterPro: IPR001827; Antennapedia.

DR InterPro: IPR001356; Homeobox.

DR Pfam: PF00046; homeobox; 1.

DR PRINTS: PR00025; ANTENNAPEDIA.

DR PRINTS: PR00024; HOMEBOX.

DR SMART: SM00389; HOX; 1.

DR PROSITE: PS00027; HOMEBOX\_1; 1.

DR PROSITE: PS00032; ANTENNAPEDIA; 1.  
DR PROSITE: PS50071; HOMEOBOX\_2; 1.  
KW Homeobox; DNA-binding; Developmental protein; Nuclear protein;  
KW Transcription regulation.  
FT DOMAIN 15 138 PRO-RICH; PART OF THE TRANSCRIPTIONAL  
FT DOMAIN 71 86 ACTIVATION DOMAIN.  
FT DOMAIN 140 145 POLY-PRO.  
FT DNA\_BIND 161 220 ANTP-TYPE HEXAPEPTIDE.  
SQ SEQUENCE 250 AA; 27519 MW; D09D477A0E585BE6 CRC64;

Query Match 1.9%; Score 7; DB 1; Length 250;  
Best Local Similarity 100.0%; Pred. No. 25;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 238 GGORRES 244  
Db 40 GGORRES 46

Search completed: August 28, 2002, 11:27:19  
Job time: 512 sec

**THIS PAGE BLANK (USPTO)**

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 28, 2002, 11:26:50 ; Search time 83.88 Seconds  
(without alignments)  
1050.387 Million cell updates/sec

Title: US-09-853-526-5  
364  
Sequence: 1 MLSTVHTYSMRILPSVY.....YVNTWIVTILGLMTVTKA 364

Scoring table: OLIGO  
Gapop 60.0 , Gapext 60.0

Searched: 747981 seqs, 242050750 residues

Word size : 0

Total number of hits satisfying chosen parameters: 747981

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : Pending\_Patents\_AA\_New:\*  
1: /cgn2\_6/ptodata/1/paa/PCT\_NEW\_COMB.pep:\*  
2: /cgn2\_6/ptodata/1/paa/US06\_NEW\_COMB.pep:\*  
3: /cgn2\_6/ptodata/1/paa/US07\_NEW\_COMB.pep:\*  
4: /cgn2\_6/ptodata/1/paa/US08\_NEW\_COMB.pep:\*  
5: /cgn2\_6/ptodata/1/paa/US09\_NEW\_COMB.pep:\*  
6: /cgn2\_6/ptodata/1/paa/US10\_NEW\_COMB.pep:\*  
7: /cgn2\_6/ptodata/1/paa/US60\_NEW\_COMB.pep:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	364	100.0	364	US-10-184-648-63	Sequence 63, Appl
2	269	73.9	269	US-10-074-045-47	Sequence 47, Appl
3	252	69.2	353	US-09-629-469A-13028	Sequence 13028, A
4	9	2.5	135	US-10-184-648-52	Sequence 52, Appl
5	9	2.5	157	US-10-184-648-66	Sequence 66, Appl
6	9	2.5	240	US-09-895-913A-184	Sequence 184, Appl
7	8	2.2	236	US-10-053-853A-1538	Sequence 1538, Ap
8	2.2	459	5	US-09-791-537-5820	Sequence 5820, Ap
9	8	2.2	459	US-09-791-537-110883	Sequence 110883,
10	8	2.2	483	US-09-791-537-110642	Sequence 110642,
11	7	1.9	44	US-09-826-733A-142	Sequence 142, Appl
12	7	1.9	77	US-09-548-936C-17	Sequence 17, Appl
13	7	1.9	96	US-09-791-537-150979	Sequence 150979,
14	7	1.9	117	US-09-791-537-31675	Sequence 31675, A
15	7	1.9	126	US-09-791-537-63528	Sequence 63528, A
16	7	1.9	194	US-09-791-537-53250	Sequence 53250, A
17	7	1.9	195	US-10-164-966-9	Sequence 9, Appl
18	7	1.9	195	US-10-184-648-51	Sequence 51, Appl
19	7	1.9	195	US-10-184-648-67	Sequence 67, Appl
20	7	1.9	195	US-10-184-648-86	Sequence 86, Appl
21	7	1.9	208	PCT-US02-10824-132	Sequence 132, Appl
22	7	1.9	208	US-09-791-537-57367	Sequence 57367, A
23	7	1.9	208	US-09-791-537-99121	Sequence 99121, A
24	7	1.9	208	US-09-791-537-128069	Sequence 128069,
25	7	1.9	208	US-09-791-537-142901	Sequence 142901,
26	7	1.9	208	US-09-791-537-143774	Sequence 143774,

27	7	1.9	208	US-10-096-327-2	Sequence 2, Appl
28	7	1.9	208	US-10-179-131-7124	Sequence 7124, Ap
29	7	1.9	208	US-10-189-360-5	Sequence 5, Appl
30	7	1.9	208	US-10-189-360-7	Sequence 7, Appl
31	7	1.9	208	US-10-189-360-8	Sequence 8, Appl
32	7	1.9	208	US-10-138-158-18	Sequence 18, Appl
33	7	1.9	234	US-09-791-537-148811	Sequence 148811,
34	7	1.9	246	US-09-791-537-43519	Sequence 43519, A
35	7	1.9	250	US-09-791-537-67119	Sequence 67119, A
36	7	1.9	251	US-09-791-537-29660	Sequence 29660, A
37	7	1.9	255	US-10-153-881-22428	Sequence 22428, A
38	7	1.9	262	US-09-791-537-751	Sequence 751, Appl
39	7	1.9	263	US-10-155-881-18654	Sequence 18654, A
40	7	1.9	264	US-09-791-537-24942	Sequence 24942, A
41	7	1.9	269	US-10-053-853A-1244	Sequence 1244, Ap
42	7	1.9	270	US-10-211-364-1026	Sequence 1026, Ap
43	7	1.9	286	US-60-360-039-1733	Sequence 1733, Ap
44	7	1.9	311	US-60-360-039-22196	Sequence 22196, A
45	7	1.9	364	US-09-791-537-42339	Sequence 42339, A

ALIGNMENTS

RESULT 1  
US-10-184-648-63  
; Sequence 63, Application US/10184648  
; GENERAL INFORMATION:  
; APPLICANT: Meyers, Rachel E.  
; APPLICANT: Williamson, Mark  
; APPLICANT: Tsai, Feng-Ying  
; APPLICANT: Hunter, John J.  
; APPLICANT: Macbeth, Kyle J.  
; APPLICANT: Rudolph-Owen, Laura A.  
; APPLICANT: Leiby, Kevin R.  
; APPLICANT: Kapeller-Libermann, Rosana  
; APPLICANT: Olandt, Peter J.  
; TITLE OF INVENTION: NOVEL HUMAN TRANSFERASE FAMILY MEMBERS AND USES THEREOF  
; FILE REFERENCE: 10448-192001  
; CURRENT APPLICATION NUMBER: US/10/184,648  
; CURRENT FILING DATE: 2002-06-27  
; PRIOR APPLICATION NUMBER: US 09/815,028  
; PRIOR FILING DATE: 2001-03-22  
; PRIOR APPLICATION NUMBER: PCT/US01/09358  
; PRIOR FILING DATE: 2001-03-22  
; PRIOR APPLICATION NUMBER: US 60/191,964  
; PRIOR FILING DATE: 2000-03-24  
; PRIOR APPLICATION NUMBER: US 09/801,220  
; PRIOR FILING DATE: 2001-03-07  
; PRIOR APPLICATION NUMBER: PCT/US01/07269  
; PRIOR FILING DATE: 2001-03-07  
; PRIOR APPLICATION NUMBER: US 60/187,456  
; PRIOR FILING DATE: 2000-03-07  
; PRIOR APPLICATION NUMBER: US 09/816,714  
; PRIOR FILING DATE: 2001-03-23  
; PRIOR APPLICATION NUMBER: PCT/US01/09468  
; PRIOR FILING DATE: 2001-03-23  
; PRIOR APPLICATION NUMBER: US 60/191,865  
; PRIOR FILING DATE: 2000-03-24  
; PRIOR APPLICATION NUMBER: US 09/844,948  
; PRIOR FILING DATE: 2001-04-27  
; PRIOR APPLICATION NUMBER: PCT/US01/13805  
; PRIOR FILING DATE: 2001-04-27  
; PRIOR APPLICATION NUMBER: US 60/200,604  
; PRIOR FILING DATE: 2000-04-28  
; PRIOR APPLICATION NUMBER: US 09/861,164  
; PRIOR FILING DATE: 2001-05-18  
; PRIOR APPLICATION NUMBER: PCT/US01/16292  
; PRIOR FILING DATE: 2001-05-18  
; PRIOR APPLICATION NUMBER: US 60/205,408  
; PRIOR FILING DATE: 2000-05-19  
; PRIOR APPLICATION NUMBER: US 09/883,060  
; PRIOR FILING DATE: 2001-06-15

```

; PRIOR APPLICATION NUMBER: PCT/US01/19138
; PRIOR FILING DATE: 2001-06-15
; PRIOR APPLICATION NUMBER: JS 60/212,079
; PRIOR FILING DATE: 2000-06-15
; PRIOR APPLICATION NUMBER: US 09/962,678
; PRIOR FILING DATE: 2001-09-25
; PRIOR APPLICATION NUMBER: PCT/US01/29963
; PRIOR FILING DATE: 2001-09-25
; PRIOR APPLICATION NUMBER: US 60/235,044
; PRIOR FILING DATE: 2000-09-25
; PRIOR APPLICATION NUMBER: US 09/973,457
; PRIOR FILING DATE: 2001-10-09
; PRIOR APPLICATION NUMBER: US 60/238,849
; PRIOR FILING DATE: 2000-10-06
; PRIOR APPLICATION NUMBER: US 10/072,285
; PRIOR FILING DATE: 2002-02-08
; PRIOR APPLICATION NUMBER: PCT/US02/03736
; PRIOR FILING DATE: 2002-02-08
; PRIOR APPLICATION NUMBER: US 60/267,494
; PRIOR FILING DATE: 2001-02-08
; PRIOR APPLICATION NUMBER: US 09/817,910
; PRIOR FILING DATE: 2001-03-16
; PRIOR APPLICATION NUMBER: PCT/US01/09633
; PRIOR FILING DATE: 2001-03-26
; PRIOR APPLICATION NUMBER: US 60/192,092
; PRIOR FILING DATE: 2000-03-24
; PRIOR APPLICATION NUMBER: US 09/842,528
; PRIOR FILING DATE: 2001-04-25
; PRIOR APPLICATION NUMBER: PCT/US01/40607
; PRIOR FILING DATE: 2001-04-25
; PRIOR APPLICATION NUMBER: US 60/199,500
; PRIOR FILING DATE: 2000-04-25
; PRIOR APPLICATION NUMBER: US 09/882,836
; PRIOR FILING DATE: 2001-06-15
; PRIOR APPLICATION NUMBER: PCT/US01/19543
; PRIOR FILING DATE: 2001-06-15
; PRIOR APPLICATION NUMBER: US 60/211,730
; PRIOR FILING DATE: 2000-06-15
; PRIOR APPLICATION NUMBER: US 09/882,872
; PRIOR FILING DATE: 2001-06-15
; PRIOR APPLICATION NUMBER: PCT/US01/19153
; PRIOR FILING DATE: 2001-06-15
; PRIOR APPLICATION NUMBER: US 50/212,077
; PRIOR FILING DATE: 2000-06-15
; NUMBER OF SEQ ID NOS: 90
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 63
; LENGTH: 364
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-184-648-63
```

```

Query Match      100.0%; Score 364; DB 6; Length 364;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 364; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```

QY 1 MLSTLVHTYSMKRYLLPSVVLGTAPTYVLAMGVWRLSAPFLPARFYQALDDRLCYCYOS 60
   |||
DB 1 MLSTLVHTYSMKRYLLPSVVLGTAPTYVLAMGVWRLSAPFLPARFYQALDDRLCYCYOS 50
QY 61 MVLFFENYTGVOILLYGDLPRKN:ENITTYLANHSTVDWIVADILAIROMALGHVRYLTK 120
   |||
DB 61 MVLFFENYTGVOILLYGDLPRKN:ENITTYLANHSTVDWIVADILAIROMALGHVRYLTK 120
QY 121 ESKKMLPLYGCTYFAOHGCTIVKRSKAFNEKEMKNKLSQSYVDAGTPTMLVLFPEGSTRNPE 180
   |||
DB 121 ESKKMLPLYGCTYFAOHGCTIVKRSKAFNEKEMKNKLSQSYVDAGTPTMLVLFPEGSTRNPE 180
QY 181 QTKVLSASQAFAGRLAVLKHVLFPRIKATHVAFDCMKRYLDAIVDYVTYVYEGKDDGGQ 240
   |||
DB 181 QTKVLSASQAFAGRLAVLKHVLFPRIKATHVAFDCMKRYLDAIVDYVTYVYEGKDDGGQ 240
```

```

QY 241 RRESPTWTEFLCKECPKIHIIHIDRDKDVPEDOEHRMRWLHREFEIKDKMLIEFESPD 300
   |||
DB 241 RRESPTWTEFLCKECPKIHIIHIDRDKDVPEDOEHRMRWLHREFEIKDKMLIEFESPD 300
QY 301 PERKKRPPGKSVNSKLSIKRTPLSMILSGLTAGMLTMDGRLKYVNTWITYGTLGCLMW 360
   |||
DB 301 PERKKRPPGKSVNSKLSIKRTPLSMILSGLTAGMLTMDGRLKYVNTWITYGTLGCLMW 360
QY 361 TIRA 364
   |||
DB 361 TIRA 364
```

```

RESULT 2
US-10-074-045-47
; Sequence 47, Application US/10074045
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PT2121C1
; CURRENT APPLICATION NUMBER: US/10/074,045
; CURRENT FILING DATE: 2002-02-14
; Prior Application removed - See file Wrapper or Palm
; NUMBER OF SEQ ID NOS: 74
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 47
; LENGTH: 269
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-074-045-47
```

```

Query Match      73.9%; Score 269; DB 6; Length 269;
Best Local Similarity 100.0%; Pred. No. 1.5e-267;
Matches 269; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```

QY 96 TVDMIVADILAIROMALGHVRYLKBGLKMLPLYGCTYFAOHGCTIVKRSKAFNEKEMKNK 155
   |||
DB 1 TVDMIVADILAIROMALGHVRYLKBGLKMLPLYGCTYFAOHGCTIVKRSKAFNEKEMKNK 155
QY 156 LQSYVDAGTPTMLVLFPEGSTRNPEOTKVLASQAFAGRLAVLKHVLFPRIKATHVAF 215
   |||
DB 61 LQSYVDAGTPTMLVLFPEGSTRNPEOTKVLASQAFAGRLAVLKHVLFPRIKATHVAF 120
QY 216 DCKKNYLDAIVDYTVYVEGKDDGQRRSEPTWTEFLCKECPKIHIIHIDRDKDVEDEE 275
   |||
DB 121 DCKKNYLDAIVDYTVYVEGKDDGQRRSEPTWTEFLCKECPKIHIIHIDRDKDVEDEE 180
QY 276 HMRMRWLHREFEIKDKMLIEFESPDPERKKRPPGKSVNSKLSIKRTPLSMILSGTAGM 335
   |||
DB 181 HMRMRWLHREFEIKDKMLIEFESPDPERKKRPPGKSVNSKLSIKRTPLSMILSGTAGM 240
QY 336 LMTDAGRKLKYVNTWITYGTLGCLMW TIRA 364
   |||
DB 241 LMTDAGRKLKYVNTWITYGTLGCLMW TIRA 269
```

```

RESULT 3
US-09-629-469A-13028
; Sequence 13028, Application US/09629469A
; GENERAL INFORMATION:
; APPLICANT: OTA, TOSHIO
; APPLICANT: ISOGAI, TAKAO
; APPLICANT: NISHIKAWA, TETSUO
; APPLICANT: HAYASHI, KOJI
; APPLICANT: SAITO, KAORI
; APPLICANT: YAMAMOTO, JUNICHI
; APPLICANT: ISHII, SHIZUKO
; APPLICANT: SUGIYAMA, TOMOYASU
; APPLICANT: WAKAMATSU, AI
; APPLICANT: NAGAI, KEIICHI
; APPLICANT: OTSUKI, TETSUJI
; TITLE OF INVENTION: PRIMERS FOR SYNTHESIZING FULL-LENGTH CDNA AND THEIR USE
```



```
FILE REFERENCE: 084335/0123
CURRENT APPLICATION NUMBER: US/09/629,469A
CURRENT FILING DATE: 2000-07-28
PRIOR APPLICATION NUMBER: JP 1999-248036
PRIOR FILING DATE: 1999-07-29
PRIOR APPLICATION NUMBER: JP 1999-300253
PRIOR FILING DATE: 1999-08-27
PRIOR APPLICATION NUMBER: JP 2000-118776
PRIOR FILING DATE: 2000-01-11
PRIOR APPLICATION NUMBER: JP 2000-183767
PRIOR FILING DATE: 2000-05-02
PRIOR APPLICATION NUMBER: JP 2000-241899
PRIOR FILING DATE: 2000-06-09
PRIOR APPLICATION NUMBER: 60/159,590
PRIOR FILING DATE: 1999-10-18
PRIOR APPLICATION NUMBER: 60/183,322
PRIOR FILING DATE: 2000-02-17
NUMBER OF SEQ ID NOS: 19025
SOFTWARE: Patent In Ver. 2.1
SEQ ID NO 13028
LENGTH: 353
TYPE: PRT
ORGANISM: Homo sapiens
US-09-629-469A-13028
```

```
Query Match          69.2%  Score 252; DB 5; Length 353;
Best Local Similarity 99.7%  Pred. No. 5.7e-250;
Matches 352; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
```

```
OY 12 MRLLPSVVLGTAPTYVVLAMGVWRLSAPLPAFYQALDDRCYCYQSQVLEFFENYTG 71
Db 1 MRLLPSVVLGTAPTYVVLAMGVWRLSAPLPAFYQALDDRCYCYQSQVLEFFENYTG 60
OY 72 VQILLGLDLPKKNENIYYLANHOSYDWIYADILAIQNALGHVRYLKEGLKWLPLYGC 131
Db 61 VQILLGLDLPKKNENIYYLANHOSYDWIYADILAIQNALGHVRYLKEGLKWLPLYGC 120
OY 132 YFQHGIGIYKRSKAKNEKEMRNKLOSVDAGTPTMYLVIPEEGTTRYPEEGTKYLSAQAF 191
Db 121 YFQHGIGIYKRSKAKNEKEMRNKLOSVDAGTPTMYLVIPEEGTTRYPEEGTKYLSAQAF 180
OY 192 AAGRGVLAVLKHVLPRIKATHVAFDCMKNYLDIYDVTVYVEGKDGQGRRESPTWTEFL 251
Db 181 AAGRGVLAVLKHVLPRIKATHVAFDCMKNYLDIYDVTVYVEGKDGQGRRESPTWTEFL 240
OY 252 CKECPIHIDRIDKKDVEEQQEHMRWMLHERFEIKDKMLIEFYESPDERRRKRFPGKS 311
Db 241 CKECPIHIDRIDKKDVEEQQEHMRWMLHERFEIKDKMLIEFYESPDERRRKRFPGKS 300
OY 312 VNSKLSIKKTLPMSMLISGLTAGMLMTDAGRKIYVNTWYIGTLLGLMTWITKA 364
Db 301 VNSKLSIKKTLPMSMLISGLTAGMLMTDAGRKIYVNTWYIGTLLGLMTWITKA 353
```

```
RESULT 4
US-10-184-648-52
Sequence 52, Application US/10184648
GENERAL INFORMATION:
APPLICANT: Meyers, Rachel E.
APPLICANT: Williamson, Mark
APPLICANT: Tsai, Feng-Ying
APPLICANT: Hunter, John J.
APPLICANT: Macbeth, Kyle J.
APPLICANT: Rudolph-Owen, Laura A.
APPLICANT: Leiby, Kevin R.
APPLICANT: Kapeller-Libermann, Rosana
APPLICANT: Olandt, Peter J.
TITLE OF INVENTION: NOVEL HUMAN TRANSFERASE FAMILY MEMBERS AND USES THEREOF
FILE REFERENCE: 10448-192001
CURRENT APPLICATION NUMBER: US/10/184,648
CURRENT FILING DATE: 2002-06-27
PRIOR APPLICATION NUMBER: US 09/815,028
```

```
PRIOR FILING DATE: 2001-03-22
PRIOR APPLICATION NUMBER: PCT/US01/09358
PRIOR FILING DATE: 2001-03-22
PRIOR APPLICATION NUMBER: US 60/191,964
PRIOR FILING DATE: 2000-03-24
PRIOR APPLICATION NUMBER: US 09/801,220
PRIOR FILING DATE: 2001-03-07
PRIOR APPLICATION NUMBER: PCT/US01/07269
PRIOR FILING DATE: 2001-03-07
PRIOR APPLICATION NUMBER: US 60/187,456
PRIOR FILING DATE: 2000-03-07
PRIOR APPLICATION NUMBER: US 09/816,714
PRIOR FILING DATE: 2001-03-23
PRIOR APPLICATION NUMBER: PCT/US01/09468
PRIOR FILING DATE: 2001-03-23
PRIOR APPLICATION NUMBER: US 60/191,865
PRIOR FILING DATE: 2000-03-24
PRIOR APPLICATION NUMBER: US 09/844,948
PRIOR FILING DATE: 2001-04-27
PRIOR APPLICATION NUMBER: PCT/US01/13805
PRIOR FILING DATE: 2001-04-27
PRIOR APPLICATION NUMBER: US 60/200,604
PRIOR FILING DATE: 2000-04-28
PRIOR APPLICATION NUMBER: US 09/861,164
PRIOR FILING DATE: 2001-05-18
PRIOR APPLICATION NUMBER: PCT/US01/16292
PRIOR FILING DATE: 2001-05-18
PRIOR APPLICATION NUMBER: US 60/205,408
PRIOR FILING DATE: 2000-05-19
PRIOR APPLICATION NUMBER: US 09/883,060
PRIOR FILING DATE: 2001-06-15
PRIOR APPLICATION NUMBER: PCT/US01/19138
PRIOR FILING DATE: 2001-06-15
PRIOR APPLICATION NUMBER: US 60/212,079
PRIOR FILING DATE: 2000-06-15
PRIOR APPLICATION NUMBER: US 09/962,678
PRIOR FILING DATE: 2001-09-25
PRIOR APPLICATION NUMBER: PCT/US01/29963
PRIOR FILING DATE: 2001-09-25
PRIOR APPLICATION NUMBER: US 60/235,044
PRIOR FILING DATE: 2000-09-25
PRIOR APPLICATION NUMBER: US 09/973,457
PRIOR FILING DATE: 2001-10-09
PRIOR APPLICATION NUMBER: US 60/238,849
PRIOR FILING DATE: 2000-10-06
PRIOR APPLICATION NUMBER: US 10/072,285
PRIOR FILING DATE: 2002-02-08
PRIOR APPLICATION NUMBER: PCT/US02/03736
PRIOR FILING DATE: 2002-02-08
PRIOR APPLICATION NUMBER: US 60/267,494
PRIOR FILING DATE: 2001-02-08
PRIOR APPLICATION NUMBER: US 09/817,910
PRIOR FILING DATE: 2001-03-26
PRIOR APPLICATION NUMBER: PCT/US01/09633
PRIOR FILING DATE: 2001-03-26
PRIOR APPLICATION NUMBER: US 60/192,092
PRIOR FILING DATE: 2000-03-24
PRIOR APPLICATION NUMBER: US 09/842,528
PRIOR FILING DATE: 2001-04-25
PRIOR APPLICATION NUMBER: PCT/US01/40607
PRIOR FILING DATE: 2001-04-25
PRIOR APPLICATION NUMBER: US 60/199,500
PRIOR FILING DATE: 2000-04-25
PRIOR APPLICATION NUMBER: US 09/882,836
PRIOR FILING DATE: 2001-06-15
PRIOR APPLICATION NUMBER: PCT/US01/19543
PRIOR FILING DATE: 2001-06-15
PRIOR APPLICATION NUMBER: US 60/211,730
PRIOR FILING DATE: 2000-06-15
PRIOR APPLICATION NUMBER: US 09/882,872
PRIOR FILING DATE: 2001-06-15
PRIOR APPLICATION NUMBER: PCT/US01/19153
PRIOR FILING DATE: 2001-06-15
```

;; PRIOR APPLICATION NUMBER: US 60/212,077  
;; PRIOR FILING DATE: 2000-06-15  
;; NUMBER OF SEQ ID NOS: 90  
;; SOFTWARE: FastSeq for Windows Version 4.0  
;; SEQ ID NO 52  
;; LENGTH: 135  
;; TYPE: PRT  
;; ORGANISM: Artificial Sequence  
;; FEATURE:  
;; OTHER INFORMATION: consensus sequence  
US-10-184-648-52

Query Match 2.5%; Score 9; DB 6; Length 135;  
Best Local Similarity 100.0%; Pred. No. 0.71;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 168 LVIFPEGTR 176  
Db 85 LVIFPEGTR 93

RESULT 5  
US-10-184-648-66  
;; Sequence 66, Application US/10184648  
;; GENERAL INFORMATION:  
;; APPLICANT: Meyers, Rache, E.  
;; APPLICANT: Williamson, Mark  
;; APPLICANT: Tsai, Fong-Ying  
;; APPLICANT: Hunter, John J.  
;; APPLICANT: Macbeth, Kyle J.  
;; APPLICANT: Rudolph-Owen, Laura A.  
;; APPLICANT: Leiby, Kevin R.  
;; APPLICANT: Kapeller-Libermann, Rosana  
;; TITLE OF INVENTION: NOVEL HUMAN TRANSFERASE FAMILY MEMBERS AND USES THEREOF  
;; FILE REFERENCE: 1048-192001  
;; CURRENT APPLICATION NUMBER: US/10/184,648  
;; PRIOR FILING DATE: 2002-06-27  
;; PRIOR APPLICATION NUMBER: US 09/815,028  
;; PRIOR FILING DATE: 2001-03-22  
;; PRIOR APPLICATION NUMBER: PCT/US01/09358  
;; PRIOR FILING DATE: 2001-03-22  
;; PRIOR APPLICATION NUMBER: US 60/191,964  
;; PRIOR FILING DATE: 2000-03-24  
;; PRIOR APPLICATION NUMBER: US 09/801,220  
;; PRIOR FILING DATE: 2001-03-07  
;; PRIOR APPLICATION NUMBER: PCT/US01/07269  
;; PRIOR FILING DATE: 2001-03-07  
;; PRIOR APPLICATION NUMBER: US 60/187,456  
;; PRIOR FILING DATE: 2000-03-07  
;; PRIOR APPLICATION NUMBER: US 09/816,714  
;; PRIOR FILING DATE: 2001-03-23  
;; PRIOR APPLICATION NUMBER: PCT/US01/09468  
;; PRIOR FILING DATE: 2001-03-23  
;; PRIOR APPLICATION NUMBER: US 60/191,865  
;; PRIOR FILING DATE: 2000-03-24  
;; PRIOR APPLICATION NUMBER: US 09/844,948  
;; PRIOR FILING DATE: 2001-04-27  
;; PRIOR APPLICATION NUMBER: PCT/US01/13805  
;; PRIOR FILING DATE: 2001-04-27  
;; PRIOR APPLICATION NUMBER: US 60/200,604  
;; PRIOR FILING DATE: 2000-04-28  
;; PRIOR APPLICATION NUMBER: US 09/861,164  
;; PRIOR FILING DATE: 2001-05-18  
;; PRIOR APPLICATION NUMBER: PCT/US01/16292  
;; PRIOR FILING DATE: 2001-05-18  
;; PRIOR APPLICATION NUMBER: US 60/205,408  
;; PRIOR FILING DATE: 2000-05-19  
;; PRIOR APPLICATION NUMBER: US 09/883,060  
;; PRIOR FILING DATE: 2001-06-15  
;; PRIOR APPLICATION NUMBER: PCT/US01/19138  
;; PRIOR FILING DATE: 2001-06-15

;; PRIOR APPLICATION NUMBER: US 60/212,079  
;; PRIOR FILING DATE: 2000-06-15  
;; PRIOR APPLICATION NUMBER: US 09/962,678  
;; PRIOR FILING DATE: 2001-09-25  
;; PRIOR APPLICATION NUMBER: PCT/US01/29963  
;; PRIOR FILING DATE: 2001-09-25  
;; PRIOR APPLICATION NUMBER: US 60/235,044  
;; PRIOR FILING DATE: 2000-09-25  
;; PRIOR APPLICATION NUMBER: US 09/973,457  
;; PRIOR FILING DATE: 2001-10-09  
;; PRIOR APPLICATION NUMBER: US 60/238,849  
;; PRIOR FILING DATE: 2000-10-06  
;; PRIOR APPLICATION NUMBER: US 10/072,285  
;; PRIOR FILING DATE: 2002-02-08  
;; PRIOR APPLICATION NUMBER: PCT/US02/03736  
;; PRIOR FILING DATE: 2002-02-08  
;; PRIOR APPLICATION NUMBER: US 60/267,494  
;; PRIOR FILING DATE: 2001-02-08  
;; PRIOR APPLICATION NUMBER: US 09/817,910  
;; PRIOR FILING DATE: 2001-03-26  
;; PRIOR APPLICATION NUMBER: PCT/US01/09633  
;; PRIOR FILING DATE: 2001-03-26  
;; PRIOR APPLICATION NUMBER: US 60/192,092  
;; PRIOR FILING DATE: 2000-03-24  
;; PRIOR APPLICATION NUMBER: US 09/842,528  
;; PRIOR FILING DATE: 2001-04-25  
;; PRIOR APPLICATION NUMBER: PCT/US01/40607  
;; PRIOR FILING DATE: 2001-04-25  
;; PRIOR APPLICATION NUMBER: US 60/199,500  
;; PRIOR FILING DATE: 2000-04-25  
;; PRIOR APPLICATION NUMBER: US 09/882,836  
;; PRIOR FILING DATE: 2001-06-15  
;; PRIOR APPLICATION NUMBER: PCT/US01/19543  
;; PRIOR FILING DATE: 2001-06-15  
;; PRIOR APPLICATION NUMBER: US 60/211,730  
;; PRIOR FILING DATE: 2000-06-15  
;; PRIOR APPLICATION NUMBER: US 09/882,872  
;; PRIOR FILING DATE: 2001-06-15  
;; PRIOR APPLICATION NUMBER: PCT/US01/19153  
;; PRIOR FILING DATE: 2001-06-15  
;; PRIOR APPLICATION NUMBER: US 60/212,077  
;; PRIOR FILING DATE: 2000-06-15  
;; NUMBER OF SEQ ID NOS: 90  
;; SOFTWARE: FastSeq for Windows Version 4.0  
;; SEQ ID NO 66  
;; LENGTH: 157  
;; TYPE: PRT  
;; ORGANISM: Artificial Sequence  
;; FEATURE:  
;; OTHER INFORMATION: consensus sequence  
US-10-184-648-66

Query Match 2.5%; Score 9; DB 6; Length 157;  
Best Local Similarity 100.0%; Pred. No. 0.81;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 222 LDATYDVT 230  
Db 4 LDATYDVT 12

RESULT 6  
US-09-895-913A-184  
;; Sequence 184, Application US/09895913A  
;; GENERAL INFORMATION:  
;; APPLICANT: Kleantous, Harold  
;; APPLICANT: Al-Garawi, Amal  
;; APPLICANT: Miller, Charles  
;; APPLICANT: Tomb, Jean Francois  
;; APPLICANT: Oomen, Raymond P.  
;; TITLE OF INVENTION: Identification of Polynucleotides  
;; TITLE OF INVENTION: Encoding Novel Helicobacter Polypeptides in the Helicobacter

```
; TITLE OF INVENTION: Genome
; FILE REFERENCE: 06132/043002
; CURRENT APPLICATION NUMBER: US/09/895,913A
; CURRENT FILING DATE: 2001-06-29
; PRIOR APPLICATION NUMBER: US 08/881,227
; PRIOR FILING DATE: 1997-06-24
; NUMBER OF SEQ ID NOS: 368
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 184
; LENGTH: 240
; TYPE: PRT
; ORGANISM: Helicobacter pylori
US-09-895-913A-184
```

```
Query Match          2.5%; Score 9; DB 5; Length 240;
Best Local Similarity 100.0%; Pred. No. 1.2;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 168 LVIFPESTR 176
      |||||
Db 142 LVIFPESTR 150
```

```
RESULT 7
US-10-053-853A-1538
; Sequence 1538, Application US/10053853A
; GENERAL INFORMATION:
; APPLICANT: HAYASHI, Hideo
; APPLICANT: SHINGAWA, Hideo
; APPLICANT: MAKINO, Kozo
; APPLICANT: HAYASHI, Tetsuya
; APPLICANT: OHNISHI, Makoto
; APPLICANT: HATTORI, Masahira
; APPLICANT: KUROKAWA, Ken
; TITLE OF INVENTION: Polynucleotide molecules and polypeptides specific to Enterohemorrhagic E. coli O157:H7 and use thereof
; FILE REFERENCE: 2002-0060A/MNC/01704
; CURRENT APPLICATION NUMBER: US/10/053,853A
; CURRENT FILING DATE: 2002-05-28
; PRIOR APPLICATION NUMBER: JP2001-112010
; PRIOR FILING DATE: 2001-01-24
; NUMBER OF SEQ ID NOS: 1866
; SEQ ID NO 1538
; LENGTH: 236
; TYPE: PRT
; ORGANISM: Escherichia coli O157:H7
US-10-053-853A-1538
```

```
Query Match          2.2%; Score 8; DB 6; Length 236;
Best Local Similarity 100.0%; Pred. No. 13;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 204 LTPRIKAT 211
      |||||
Db 227 LTPRIKAT 234
```

```
RESULT 8
US-09-791-537-5820
; Sequence 5820, Application US/09791537
; GENERAL INFORMATION:
; APPLICANT: Bionomix, Inc.
; APPLICANT: Dabez, Derek
; APPLICANT: Danzer, Joseph
; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMBERS
; FILE REFERENCE: 261/210
; CURRENT APPLICATION NUMBER: US/09/791,537
; CURRENT FILING DATE: 2001-02-22
; NUMBER OF SEQ ID NOS: 153055
; SOFTWARE: PatentIn version 3.0
```

```
; SEQ ID NO 5820
; LENGTH: 262
; TYPE: PRT
; ORGANISM: Chlamydia trachomatis
US-09-791-537-5820
```

```
Query Match          2.2%; Score 8; DB 5; Length 262;
Best Local Similarity 100.0%; Pred. No. 14;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 202 HVTPESTR 209
      |||||
Db 202 HVTPESTR 209
```

```
RESULT 9
US-09-791-537-110883
; Sequence 110883, Application US/09791537
; GENERAL INFORMATION:
; APPLICANT: Bionomix, Inc.
; APPLICANT: Dabez, Derek
; APPLICANT: Danzer, Joseph
; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMBERS
; FILE REFERENCE: 261/210
; CURRENT APPLICATION NUMBER: US/09/791,537
; CURRENT FILING DATE: 2001-02-22
; NUMBER OF SEQ ID NOS: 153055
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 110883
; LENGTH: 459
; TYPE: PRT
; ORGANISM: Bacillus subtilis
US-09-791-537-110883
```

```
Query Match          2.2%; Score 8; DB 5; Length 459;
Best Local Similarity 100.0%; Pred. No. 24;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 178 NPEQTKVL 185
      |||||
Db 292 NPEQTKVL 299
```

```
RESULT 10
US-09-791-537-110642
; Sequence 110642, Application US/09791537
; GENERAL INFORMATION:
; APPLICANT: Bionomix, Inc.
; APPLICANT: Dabez, Derek
; APPLICANT: Danzer, Joseph
; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMBERS
; FILE REFERENCE: 261/210
; CURRENT APPLICATION NUMBER: US/09/791,537
; CURRENT FILING DATE: 2001-02-22
; NUMBER OF SEQ ID NOS: 153055
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 110642
; LENGTH: 483
; TYPE: PRT
; ORGANISM: Pium sativum
US-09-791-537-110642
```

```
Query Match          2.2%; Score 8; DB 5; Length 483;
Best Local Similarity 100.0%; Pred. No. 25;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 162 AGTPMYLV 169
      |||||
```

Db 184 AGTPMYLV 191

RESULT 11

US-09-826-734A-142

; Sequence 142, Application US/09826734A

; GENERAL INFORMATION:

; APPLICANT: Fernandes, Wilma R

; APPLICANT: Mishra, Vishnu S

; APPLICANT: Leach, Martin D

; APPLICANT: Shlmkets, Richard A

; APPLICANT: Zernhusen, Bryan D

; TITLE OF INVENTION: Novel Polynucleotides And Polypeptides Encoded Thereby

; FILE REFERENCE: 15966-754

; CURRENT APPLICATION NUMBER: US/09/826,734A

; PRIOR FILING DATE: 2001-04-05

; PRIOR APPLICATION NUMBER: 60/195,576

; PRIOR FILING DATE: 2000-04-26

; NUMBER OF SEQ ID NOS: 264

; SOFTWARE: PatentIn Ver. 2.1

; SEQ ID NO 142

; LENGTH: 44

; TYPE: PRT

; ORGANISM: Homo sapiens

US-09-826-734A-142

Query Match

Best Local Similarity 1.9%; Score 7; DB 5; Length 44;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 170 IFEPTGR 176

Db 3 IFEPTGR 9

RESULT 12

US-09-548-936C-17

; Sequence 17, Application US/09548936C

; GENERAL INFORMATION:

; APPLICANT: Gopez, Leonel Jorge

; APPLICANT: Claesson-Welsh, Lena

; APPLICANT: Heldin, Carl-Henrik

; TITLE OF INVENTION: PTPL BINDING AGENTS

; FILE REFERENCE: L0461/7084

; CURRENT APPLICATION NUMBER: US/09/548,936C

; CURRENT FILING DATE: 2000-04-13

; PRIOR APPLICATION NUMBER: US 09/100,804

; PRIOR FILING DATE: 1998-06-19

; PRIOR APPLICATION NUMBER: US 08/596,291

; PRIOR FILING DATE: 1994-09-01

; PRIOR APPLICATION NUMBER: US 08/115,573

; PRIOR FILING DATE: 1993-09-01

; NUMBER OF SEQ ID NOS: 35

; SOFTWARE: PatentIn Version 3.0

; SEQ ID NO 17

; LENGTH: 77

; TYPE: PRT

; ORGANISM: Homo sapiens

US-09-548-936C-17

Query Match

Best Local Similarity 1.9%; Score 7; DB 5; Length 77;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 136 HGGIYVK 142

Db 20 HGGIYVK 26

RESULT 13

US-09-791-537-150979

; Sequence 150979, Application US/09791537

; GENERAL INFORMATION:

; APPLICANT: Biomomix, Inc.

; APPLICANT: Debe, Derek

; APPLICANT: Danzer, Joseph

; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY ME

; FILE REFERENCE: 261/210

; CURRENT APPLICATION NUMBER: US/09/791,537

; CURRENT FILING DATE: 2001-02-22

; NUMBER OF SEQ ID NOS: 153055

; SOFTWARE: PatentIn Version 3.0

; SEQ ID NO 150979

; LENGTH: 96

; TYPE: PRT

; ORGANISM: pdb 3PDZA

; FEATURE:

; NAME/KEY: misc.feature

; LOCATION: (1)..(1)

; OTHER INFORMATION: X is an unknown amino acid

US-09-791-537-150979

Query Match

Best Local Similarity 1.9%; Score 7; DB 5; Length 96;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 136 HGGIYVK 142

Db 32 HGGIYVK 38

RESULT 14

US-09-791-537-31675

; Sequence 31675, Application US/09791537

; GENERAL INFORMATION:

; APPLICANT: Biomomix, Inc.

; APPLICANT: Debe, Derek

; APPLICANT: Danzer, Joseph

; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY ME

; FILE REFERENCE: 261/210

; CURRENT APPLICATION NUMBER: US/09/791,537

; CURRENT FILING DATE: 2001-02-22

; NUMBER OF SEQ ID NOS: 153055

; SOFTWARE: PatentIn Version 3.0

; SEQ ID NO 31675

; LENGTH: 117

; TYPE: PRT

; ORGANISM: Mus musculus

US-09-791-537-31675

Query Match

Best Local Similarity 1.9%; Score 7; DB 5; Length 117;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 136 HGGIYVK 142

Db 43 HGGIYVK 49

RESULT 15

US-09-791-537-63528

; Sequence 63528, Application US/09791537

; GENERAL INFORMATION:

; APPLICANT: Biomomix, Inc.

; APPLICANT: Debe, Derek

; APPLICANT: Danzer, Joseph

; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY ME

; METHODS OF USE THEREOF

; FILE REFERENCE: 261/210  
; CURRENT APPLICATION NUMBER: US/09/791,537  
; CURRENT FILING DATE: 2001-02-22  
; NUMBER OF SEQ ID NOS: 153055  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 63528  
; LENGTH: 126  
; TYPE: PRT  
; ORGANISM: Mus musculus  
US-09-791-537-63528

Query Match 1.9%; Score 7; DB 5; Length 126;  
Best Local Similarity 100.0%; Pred. No. 75;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 136 HGGIYVK 142  
Db 52 HGGIYVK 58

Search completed: August 28, 2002, 11:26:50  
Job time: 563 sec

**THIS PAGE BLANK (uspto)**

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 28, 2002, 11:20:08 ; Search time 39.16 Seconds  
(without alignments)  
893.169 Million cell updates/sec

Title: US-09-853-526-5  
364  
Sequence: 1 MLSLVLFHTYSMRYLPSVY.....YVNTWIVGTLLGLMTWITKA 364

Scoring table: OLIGO  
Gapop 60.0 , Gapext 60.0

Searched: 283138 seqs, 96089334 residues

Word size : 0

Total number of hits satisfying chosen parameters: 283138

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database :  
1: PIR\_71:\*  
2: pir1:\*  
3: pir2:\*  
4: pir3:\*  
5: pir4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	9	2.5	237	B71827	probable 1-acylgly
2	9	2.5	240	D64688	probable 1-acylgly
3	9	2.5	264	E82903	1-acyl-sn-glycerol
4	9	2.5	918	T34057	hypothetical prote
5	8	2.2	83	W8BP67	gene 18.7 protein
6	8	2.2	83	W8BP67	gene 18.7 protein
7	8	2.2	236	H90892	probable fibritial
8	8	2.2	236	B87524	1-acyl-sn-glycerol
9	8	2.2	239	A85725	probable metal dep
10	8	2.2	262	A71478	probable metal dep
11	8	2.2	282	G71334	probable lysophosp
12	8	2.2	396	S54641	probable membrane
13	8	2.2	411	C87586	metal ion efflux m
14	8	2.2	424	T49607	guanine deaminase
15	8	2.2	458	G83735	RNA methyltransfer
16	8	2.2	459	E69793	62K sucrose-bindin
17	8	2.2	483	T06459	heat shock protein
18	8	2.2	611	A85000	heparin-binding Eg
19	8	2.2	54	S58122	hypothetical prote
20	7	1.9	57	D86108	heparin-binding Eg
21	7	1.9	117	I81209	lysine phosphata
22	7	1.9	126	I81210	lysine phosphata
23	7	1.9	141	S58117	EGF-like growth fa
24	7	1.9	143	B90523	hypothetical prote
25	7	1.9	148	C86732	transcription regu
26	7	1.9	194	PC1136	amitophosphoribosy
27	7	1.9	195	C83854	hypothetical prote
28	7	1.9	208	A38432	heparin-binding Eg
29	7	1.9	208	A41914	diphtheria toxin re

30	7	1.9	208	1	JC1409	heparin-binding Eg
31	7	1.9	209	2	B83329	probable acyltrans
32	7	1.9	210	2	A10342	probable acyltrans
33	7	1.9	211	2	E70476	2-acylglycerophosp
34	7	1.9	225	2	H95244	conserved hypothet
35	7	1.9	226	2	E98109	conserved hypothet
36	7	1.9	230	2	E81397	probable 1-acylgly
37	7	1.9	240	2	S75162	hypothetical prote
38	7	1.9	241	2	B97019	1-acyl-sn-glycerol
39	7	1.9	246	2	T46446	hypothetical prote
40	7	1.9	247	2	A81957	1-acylglycerol-3-p
41	7	1.9	247	2	G81013	1-acyl-sn-glycerol
42	7	1.9	250	1	A31757	homeotic protein H
43	7	1.9	250	2	T27772	26S proteasome reg
44	7	1.9	250	2	E70104	1-acylglycerol-3-p
45	7	1.9	250	2	AF3384	outer membrane pro

ALIGNMENTS

RESULT 1  
B71827  
Probable 1-acylglycerol-3-phosphate O-acyltransferase (EC 2.3.1.51) - Helicobacter pylori  
C:Species: Helicobacter pylori  
A:Variate: strain J99  
C>Date: 12-Feb-1999 #sequence\_revision 12-Feb-1999 #text\_change 06-Oct-2000  
C:Accession: B71827  
R:Alm, R.A.; Ling, L.S.L.; Moir, D.T.; King, B.L.; Brown, E.D.; Doig, P.C.; Smith, D.; Ives, C.; Gibson, R.; Merberg, D.; Mills, S.D.; Jiang, O.; Taylor, D.E.; Vovis, G.F. Nature 397, 176-180, 1999  
A>Title: Genomic sequence comparison of two unrelated isolates of the human gastric p  
A:Reference number: A71800; MUID:99120557  
A:Accession: B71827  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-237 <ARN>  
A:Cross-references: GB:AE001550; GB:AE001439; NID:g4155872; PIDN:AAD06852.1; PID:g415  
A:Experimental source: strain J99  
C:Genetics:  
A:Gene: p1sc  
C:Superfamily: mouse 1-acylglycerol-3-phosphate O-acyltransferase  
C:Keywords: acyltransferase; coenzyme A

Query Match 2.5%; Score 9; DB 2; Length 237;  
Best Local Similarity 100.0%; Pred. No. 0.38;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 168 LVTFPEGTR 176  
DB 142 LVTFPEGTR 150

RESULT 2  
D64688  
Probable 1-acylglycerol-3-phosphate O-acyltransferase (EC 2.3.1.51) - Helicobacter pylori  
C:Species: Helicobacter pylori  
C>Date: 09-Aug-1997 #sequence\_revision 09-Aug-1997 #text\_change 06-Oct-2000  
C:Accession: D64688  
R:Tomp, J.F.; White, O.; Kerlavage, A.R.; Clayton, R.A.; Sutton, G.G.; Fleischmann, R.; Peterson, S.; Loftus, B.; Richardson, D.; Dodson, R.; Khajal, H.G.; Glodek, A.; McKee, J.D.; Kelley, J.M.; Cotton, M.D.; Weidman, J.M.; Fujii, C.; Bowman, C.; Wattey, Nature 388, 539-547, 1997  
A:Authors: Wallin, E.; Hayes, W.S.; Borodovsky, M.; Karpk, P.D.; Smith, H.O.; Fraser, A.; Title: The complete genome sequence of the gastric pathogen Helicobacter pylori.  
A:Reference number: A64520; MUID:97394467  
A:Accession: D64688  
A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-240 <TOM>  
A:Cross-references: GB:AE000636; GB:AE000511; NID:g2314517; PIDN:AAD08393.1; PID:g231  
C:Superfamily: mouse 1-acylglycerol-3-phosphate O-acyltransferase

C;Keywords: acyltransferase; coenzyme A

Query Match 2.5%; Score 9; DB 2; Length 240;  
Best Local Similarity 100.0%; Pred. No. 0.39;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 168 LVIFPECTR 176  
Db 142 LVIFPECTR 150

RESULT 3  
E82903  
1-acyl-sn-glycerol-3-phosphate acyltransferase U0344 [Imported] - Ureaplasma urealyticum  
C;Species: Ureaplasma urealyticum  
C;Date: 18-Aug-2000 #sequence\_revision 20-Aug-2000 #text\_change 20-Aug-2000  
C;Accession: E82903  
R;Glass, J.I.; Lefkowitz, E.J.; Glass, J.S.; Heiner, C.R.; Chen, E.Y.; Cassell, G.H.  
submitted to GenBank, February 2000  
A;Description: The complete sequence of Ureaplasma urealyticum: Alternate views of a mit  
A;Reference number: A82870  
A;Accession: E82903  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-264 <GLA>  
A;Cross-references: GB:AE002131; GB:AF222894; NID:96899316; PIDN:AAF30753.1; GSPDB:GN001  
C;Experimental source: serovar 3; biovar 1  
C;Genetics:  
A;Gene: pLSC; U0344  
A;Genetic code: SGC3

Query Match 2.5%; Score 9; DB 2; Length 264;  
Best Local Similarity 100.0%; Pred. No. 0.42;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 168 LVIFPECTR 176  
Db 156 LVIFPECTR 164

RESULT 4  
T34057  
hypothetical protein F28B3.5 - Caenorhabditis elegans  
C;Species: Caenorhabditis elegans  
C;Date: 29-Oct-1999 #sequence\_revision 29-Oct-1999 #text\_change 29-Oct-1999  
C;Accession: T34057  
R;Geisler, C.; Kramer, J.; Smith, A.  
submitted to the EMBL Data Library, May 1997  
A;Description: The sequence of C. elegans cosmid F28B3.  
A;Reference number: Z21469  
A;Accession: T34057  
A;Status: preliminary; translated from GB/EMBL/DBD  
A;Molecule type: DNA  
A;Residues: 1-918 <GET>  
A;Cross-references: EMBL:AF00136; PIDN:AA83636.1; GSPDB:GN00019; CESP:F28B3.5  
A;Experimental source: strain Bristol N2; clone F28B3  
C;Genetics:  
A;Gene: CESP:F28B3.5  
A;Map position: 1  
A;Introns: 85/3; 129/3; 235/3; 411/2; 482/3; 532/3; 736/3; 829/3; 850/2

Query Match 2.5%; Score 9; DB 2; Length 918;  
Best Local Similarity 100.0%; Pred. No. 1.3;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 222 LDAIYDVT 230  
Db 762 LDAIYDVT 770

RESULT 5  
W8BP7  
gene 18.7 protein - phage T7  
C;Species: phage T7  
C;Date: 13-Jun-1983 #sequence\_revision 13-Jun-1983 #text\_change 23-Jul-1999  
C;Accession: A04413; S42337  
R;Dunn, J.C.; Thompson, K.  
submitted to the Nucleic Acid Sequence Database, September 1982  
A;Reference number: A94615  
A;Accession: A04413  
A;Molecule type: DNA  
A;Residues: 1-83 <DUN>  
R;Dunn, J.C.; Studier, F.W.  
J. Mol. Biol. 166, 477-535, 1983  
A;Title: Complete nucleotide sequence of bacteriophage T7 DNA and the locations of T7  
A;Reference number: S42283; MUID:83241725  
A;Accession: S42337  
A;Molecule type: DNA  
A;Residues: 1-83 <DUN>  
A;Cross-references: EMBL:V01146; NID:9431187; PIDN:CAA24439.1; PID:915615  
A;Note: the authors did not translate the codon for residue 1  
C;Genetics:  
A;Gene: 18.7  
A;Map position: 92.73-93.35  
C;Superfamily: phage T7 gene 18.7 protein

Query Match 2.2%; Score 8; DB 1; Length 83;  
Best Local Similarity 100.0%; Pred. No. 1.6;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 316 LSIKKTLP 323  
Db 22 LSIKKTLP 29

RESULT 6  
W8BP73  
gene 18.7 protein - phage T3  
C;Species: phage T3  
A;Note: host Escherichia coli  
C;Date: 31-Dec-1988 #sequence\_revision 31-Dec-1988 #text\_change 23-Jul-1999  
C;Accession: E23476  
R;Yamada, M.; Fujisawa, H.; Kato, H.; Hamada, K.; Minagawa, T.  
Virology 151, 350-361, 1986  
A;Title: Cloning and sequencing of the genetic right end of bacteriophage T3 DNA.  
A;Reference number: A94339; MUID:86209997  
A;Accession: E23476  
A;Molecule type: DNA  
A;Residues: 1-83 <YAM>  
A;Cross-references: GB:M14784; NID:9215810; PIDN:AA92527.1; PID:91196765  
R;Yamada, M.; Fujisawa, H.; Kato, H.; Hamada, K.; Minagawa, T.  
Virology 154, 246, 1986  
A;Reference number: A94344  
A;Contents: annotation; erratum; corrections to coding regions  
C;Genetics:  
A;Gene: 18.7  
C;Superfamily: phage T7 gene 18.7 protein

Query Match 2.2%; Score 8; DB 1; Length 83;  
Best Local Similarity 100.0%; Pred. No. 1.6;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 316 LSIKKTLP 323  
Db 22 LSIKKTLP 29

RESULT 7  
H90892  
probable fimbrial chapone protein precursor [Imported] - Escherichia coli (strain O  
C;Species: Escherichia coli



C.Date: 18-Jul-2001 #sequence\_revision 18-Jul-2001 #text\_change 03-Aug-2001  
C.Accession: H90892  
R.Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G.  
ogawara, N.; Yatsunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shingawa, H.  
DNA Res. 8, 11-22, 2001  
A.Title: Complete genome sequence of enterohemorrhagic *Escherichia coli* O157:H7 and geno  
A.Reference number: A99629; MUID:21156231; PMID:11258796  
A.Accession: H90892  
A.Status: preliminary  
A.Molecule type: DNA  
A.Residues: 1-236 <HAN>  
A.Cross-references: GB:BA000007; PIDN:BA835535.1; PID:g13361578; GSPDB:GN00154  
A.Experimental source: strain O157:H7, substrain RMD 0509952  
C.Genetics:  
A.Gene: ECs2112  
C.Superfamily: chaperone protein papp

Query Match 2.2%; Score 8; DB 2; Length 236;  
Best Local Similarity 100.0%; Pred. No. 4.2;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 204 LTPRIKAT 211  
|||||||  
Db 227 LTPRIKAT 234

RESULT 8  
B87524  
1-acyl-sn-glycerol-3-phosphate acyltransferase [Imported] - *Caulobacter crescentus*  
C.Species: *Caulobacter crescentus*  
C.Date: 20-Apr-2001 #sequence\_revision 20-Apr-2001 #text\_change 20-Apr-2001  
C.Accession: B87524  
R.Nierman, W.C.; Feldblyum, T.V.; Paulsen, I.T.; Nelson, K.E.; Eisen, J.; Heidelberg, J.  
B.; Land, M.T.; DeBoy, R.T.; Dodson, R.J.; Durkin, A.S.; Gwinn, M.L.; Haft, D.H.; Kolon  
n, J.; Esmolaeva, M.; White, O.; Salzberg, S.L.; Shapiro, L.; Venter, J.C.; Fraser, C.M.  
Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001  
A.Title: Complete Genome Sequence of *Caulobacter crescentus*.  
A.Reference number: A87249; MUID:21173698; PMID:11259647  
A.Accession: B87524  
A.Status: preliminary  
A.Molecule type: DNA  
A.Residues: 1-236 <STO>  
A.Cross-references: GB:AE005673; NID:g13423724; PIDN:AAK24190.1; GSPDB:GN00148  
C.Genetics:  
A.Gene: CC2219

Query Match 2.2%; Score 8; DB 2; Length 236;  
Best Local Similarity 100.0%; Pred. No. 4.2;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 169 VIFPECTR 176  
|||||||  
Db 146 VIFPECTR 153

RESULT 9  
A85725  
probable fimbrial chaperone protein z2201 [Imported] - *Escherichia coli* (strain O157:H7,  
C.Species: *Escherichia coli*  
C.Date: 16-Feb-2001 #sequence\_revision 16-Feb-2001 #text\_change 14-Sep-2001  
C.Accession: A85725  
R.Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhew  
iller, L.; Grothbeck, E.J.; Davis, N.W.; Lam, A.; Dimalanta, E.; Potamousis, K.; Apodaca,  
Nature 409, 529-533, 2001  
A.Title: Genome sequence of enterohemorrhagic *Escherichia coli* O157:H7.  
A.Reference number: A85480; MUID:21074935; PMID:11206551  
A.Accession: A85725  
A.Status: preliminary  
A.Molecule type: DNA  
A.Residues: 1-239 <STO>  
A.Cross-references: GB:AE005174; NID:g12515164; PIDN:AA656261.1; GSPDB:GN00145; UMGF:222

A.Experimental source: strain O157:H7, substrain EDL933  
C.Genetics:  
A.Gene: Z2201  
C.Superfamily: chaperone protein papp

Query Match 2.2%; Score 8; DB 2; Length 239;  
Best Local Similarity 100.0%; Pred. No. 4.3;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 204 LTPRIKAT 211  
|||||||  
Db 230 LTPRIKAT 237

RESULT 10  
A71478  
probable metal dependent hydrolase - *Chlamydia trachomatis* (serotype D, strain UW3/Cx  
C.Species: *Chlamydia trachomatis*  
C.Date: 13-Sep-1998 #sequence\_revision 13-Sep-1998 #text\_change 08-Oct-1999  
C.Accession: A71478  
R.Stephens, R.S.; Kalman, S.; Lammel, C.J.; Fan, J.; Marathe, R.; Aravind, L.; Mitche  
Science 282, 754-759, 1998  
A.Title: Genome sequence of an obligate intracellular pathogen of humans: *Chlamydia t*  
A.Reference number: A71570; MUID:99000809  
A.Accession: A71478  
A.Status: preliminary  
A.Molecule type: DNA  
A.Residues: 1-262 <ARN>  
A.Cross-references: GB:AE001344; GB:AE001273; NID:g3329188; PIDN:AA68333.1; PID:g332  
A.Experimental source: serotype D, strain UW-3/Cx  
C.Genetics:  
A.Gene: yycJ

Query Match 2.2%; Score 8; DB 2; Length 262;  
Best Local Similarity 100.0%; Pred. No. 4.6;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 202 HVLTPRIK 209  
|||||||  
Db 202 HVLTPRIK 209

RESULT 11  
G71334  
probable lysophosphatidic acid acyltransferase - *Syphilis spirochete*  
C.Species: *Treponema pallidum* subsp. *pallidum* (*Syphilis spirochete*)  
C.Date: 24-Jul-1998 #sequence\_revision 24-Jul-1998 #text\_change 05-Nov-1999  
C.Accession: G71334  
R.Fraser, C.M.; Norris, S.J.; Weinstock, G.M.; White, O.; Sutton, G.G.; Dodson, R.; G  
rson, J.; Khalak, H.; Richardson, D.; Howell, J.R.; Chidambaram, M.; Uitterback, T.; M  
they, L.; Weidman, J.; Smith, H.O.; Venter, J.C.  
Science 281, 375-388, 1998  
A.Title: Complete genome sequence of *Treponema pallidum*, the *Syphilis spirochete*.  
A.Reference number: A71250; MUID:98332770  
A.Accession: G71334  
A.Status: preliminary; nucleic acid sequence not shown; translation not shown  
A.Molecule type: DNA  
A.Residues: 1-282 <COL>  
A.Cross-references: GB:AE001215; GB:AE000520; NID:g3322631; PIDN:AA65346.1; PID:g332  
A.Experimental source: strain Nichols  
C.Genetics:  
A.Gene: TP0361

Query Match 2.2%; Score 8; DB 2; Length 282;  
Best Local Similarity 100.0%; Pred. No. 5;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 169 VIFPECTR 176  
|||||||  
Db 170 VIFPECTR 177

## RESULT 12

S54641

probable membrane protein YDR018c - yeast (*Saccharomyces cerevisiae*)

N/Alternate names: hypothetical protein D3246; hypothetical protein P2F396; hypothetical

C/Species: *Saccharomyces cerevisiae*

C/Date: 08-Jul-1995 #sequence\_revision 01-Sep-1995 #text\_change 20-Jun-2000

C/Accession: S54641; S63425; S67831; S72116

R/Editor: K.J. Brown, D.J. Hamlyn, N.J. Bowman, S.

submitted to the EMBL Data Library, May 1995

A/Reference number: S54638

A/Accession: S54641

A/Molecule type: DNA

A/Residues: 1-396 &lt;DED&gt;

A/Cross-references: EMBL:Z49770; NID:q840867; PIDN:CAA9843.1; PID:q840871

A/Experimental source: strain AB972

R/Editor: L.G. Sander, C.J. Prydz, H.

submitted to the EMBL Data Library, February 1996

A/Description: Sequencing and analysis of a 35.4 kb region on the left arm of chromosome

A/Reference number: S63416

A/Accession: S63425

A/Molecule type: DNA

A/Residues: 1-396 &lt;EID&gt;

A/Cross-references: EMBL:X95966; NID:q1216215; PIDN:CAA65210.1; PID:q1216225

R/Editor: H.J. Elde, L.G.

submitted to the Protein Sequence Database, July 1996

A/Reference number: S67822

A/Accession: S67831

A/Molecule type: DNA

A/Residues: 1-396 &lt;PRY&gt;

A/Cross-references: EMBL:Z74314; NID:q1431443; PIDN:CAA9838.1; PID:q1431444; MIPS:YDR01

A/Experimental source: strain S288C

R/Editor: L.G. Sander, C.J. Prydz, H.

Yeast 12, 1085-1090, 1996

A/Title: Sequencing and analysis of a 35.4 kb region on the left arm of chromosome IV fr

A/Reference number: S72107; M01D:97051598

A/Accession: S72116

A/Molecule type: DNA

A/Residues: 1-396 &lt;ETW&gt;

A/Cross-references: EMBL:X95966; NID:q1216215; PIDN:CAA65210.1; PID:q1216225

A/Note: the nucleotide sequence was submitted to the EMBL Data Library, February 1996

C/Genetics:

A/Map position: 4R

A/Note: YDR018c

C/Superfamily: probable membrane protein YBR042c

C/Keywords: transmembrane protein

F:27-43/Domain: transmembrane #status predicted &lt;TM1&gt;

F:69-85/Domain: transmembrane #status predicted &lt;TM2&gt;

F:376-392/Domain: transmembrane #status predicted &lt;TM3&gt;

Query Match 2.2% Score 8; DB 2; Length 396;

Best Local Similarity 100.0%; Pred. No. 6.8;

Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 222 LDAIYDVT 229

Db 265 LDAIYDVT 272

## RESULT 13

C87586

metal ion efflux membrane fusion protein family [Imported] - *Caulobacter crescentus*C/Species: *Caulobacter crescentus*

C/Date: 20-Apr-2001 #sequence\_revision 20-Apr-2001 #text\_change 30-Jun-2001

C/Accession: C87586

R/Editor: W.C. Feldblyum, T.V. Paulsen, I.T. Nelson, K.E. Eisen, J.J. Heidelberg, J.

B. Laub, M.T. Deboy, R.T. Dodson, R.J. Durkin, A.S. Gwynn, M.L. Haft, D.H. Kolod

n, J.J. Ermolenko, M. White, O. Salberg, S.L. Shapiro, L. Venter, J.C. Fraser, C.M.

Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001

A/Title: Complete Genome Sequence of *Caulobacter crescentus*.

A/Reference number: A87249; M01D:21173698; PMID:11259647

A/Accession: C87586

A/Status: Preliminary

A/Molecule type: DNA

A/Residues: 1-411 &lt;STO&gt;

A/Cross-references: GB:AE005673; NID:q13424311; PIDN:AAK24687.1; GSPDB:GN00148

C/Genetics:

C/Superfamily: nickel-cobalt resistance determinant structural protein CnrB; lipoY1/b

Query Match 2.2% Score 8; DB 2; Length 411;

Best Local Similarity 100.0%; Pred. No. 7;

Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 327 ILSGITAG 334

Db 380 ILSGITAG 387

## RESULT 14

T49607

guanine deaminase (Gda) related protein [Imported] - *Neurospora crassa*

N/Alternate names: protein B3E4.200

C/Species: *Neurospora crassa*

C/Date: 02-Jun-2000 #sequence\_revision 02-Jun-2000 #text\_change 02-Jun-2000

C/Accession: T49607

R/Schulte, U.; Aign, V.; Hohenseil, J.; Brandt, P.; Fartmann, B.; Holland, R.; Nyakatu

submitted to the Protein Sequence Database, May 2000

A/Reference number: Z25022

A/Accession: T49607

A/Status: Preliminary

A/Molecule type: DNA

A/Residues: 1-424 &lt;SCH&gt;

A/Cross-references: EMBL:AL355931; GSPDB:GN00116; NCSP:B3E4.200

A/Experimental source: BAC clone B3E4; strain OR74A

C/Genetics:

A/Gene: NCSP:B3E4.200

A/Map position: 6

Query Match 2.2% Score 8; DB 2; Length 424;

Best Local Similarity 100.0%; Pred. No. 7.2;

Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 200 LKHVLPFR 207

Db 178 LKHVLPFR 185

## RESULT 15

G83735

RNA methyltransferase BH0687 [Imported] - *Bacillus halodurans* (strain C-125)C/Species: *Bacillus halodurans*

C/Date: 01-Dec-2000 #sequence\_revision 01-Dec-2000 #text\_change 15-Jun-2001

C/Accession: G83735

R/Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Masui, N.; Fujii, F.; H

Nucleic Acids Res. 28, 4317-4331, 2000

A/Title: Complete genome sequence of the alkaliphilic bacterium *Bacillus halodurans* a

A/Reference number: A83650; M01D:20512582; PMID:11058132

A/Accession: G83735

A/Status: Preliminary

A/Molecule type: DNA

A/Residues: 1-458 &lt;STO&gt;

A/Cross-references: GB:AP001509; GB:BA000004; NID:q10173176; PIDN:BA804406.1; GSPDB:G

A/Experimental source: strain C-125

C/Genetics:

A/Gene: BH0687

Query Match 2.2% Score 8; DB 2; Length 458;

Best Local Similarity 100.0%; Pred. No. 7.8;

Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy	178	NPEQTKVL	185
Db	291	NPEQTKVL	298

Search completed: August 28, 2002, 11:20:09  
Job time: 262 sec

**THIS PAGE BLANK (USPTO)**

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 28, 2002, 11:28:34 ; Search time 67.07 seconds  
(without alignments)  
586.085 Million cell updates/sec

Title: US-09-853-526-70

Perfect score: 228

Sequence: 1 MRYLPSVVLGTAPTYVLA.....NYLDATYDVTVVEGRKDDG 228

Scoring table:

Gapop 60.0 , Gapext 60.0

Searched: 562222 segs, 172994929 residues

Word size: 0

Total number of hits satisfying chosen parameters: 562222

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database:

SPTREMBL\_19:\*

- 1: sp\_archaea:\*
- 2: sp\_bacteria:\*
- 3: sp\_fungi:\*
- 4: sp\_human:\*
- 5: sp\_invertebrate:\*
- 6: sp\_mammal:\*
- 7: sp\_mhc:\*
- 8: sp\_organelle:\*
- 9: sp\_phage:\*
- 10: sp\_plant:\*
- 11: sp\_rodent:\*
- 12: sp\_virus:\*
- 13: sp\_unclassified:\*
- 14: sp\_unclassified:\*
- 15: sp\_virus:\*
- 16: sp\_bacteriap:\*
- 17: sp\_archaeap:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query	Length	DB	ID	Description
1	228	100.0	364	4	Q9B0G4	Q9B0G4 homo sapien
2	144	63.2	353	4	Q9N0U2	Q9N0U2 homo sapien
3	32	14.0	354	11	Q9D1E8	Q9D1E8 mus musculu
4	9	3.9	264	16	Q9P0E7	Q9P0E7 ureaplasma
5	5	3.9	386	5	Q95R12	Q95R12 caenobabdi
6	3	3.5	115	16	Q98E87	Q98E87 rhizobium l
7	3	3.5	236	16	Q9A675	Q9A675 caulobacter
8	3	3.5	262	16	Q84743	Q84743 chlamydia t
9	3	3.5	282	16	Q83380	Q83380 treponema p
10	3	3.5	454	3	Q9P5W2	Q9P5W2 neuropeptid
11	3	3.5	458	16	Q9KF10	Q9KF10 bacillus su
12	3	3.5	459	16	Q31503	Q31503 bacillus su
13	3	3.5	483	10	Q49927	Q49927 pisum sativ
14	8	3.5	1820	5	Q9VVG0	Q9VVG0 drosophila
15	7	3.1	54	6	Q28219	Q28219 cercopithec
16	7	3.1	126	11	Q62370	Q62370 mus musculu

17	7	3.1	132	11	Q9D581	Q9D581 mus musculu
18	7	3.1	137	12	Q55581	Q55581 leucania se
19	7	3.1	141	6	Q28218	Q28218 cercopithec
20	7	3.1	143	16	Q98R89	Q98R89 mycoplasma
21	7	3.1	148	16	Q9CH79	Q9CH79 lactococcus
22	7	3.1	150	2	Q9AGV0	Q9AGV0 bruceella ab
23	7	3.1	162	10	Q9LPP0	Q9LPP0 arabidopsis
24	7	3.1	163	2	Q9EX02	Q9EX02 streptomyces
25	7	3.1	175	2	Q9EX12	Q9EX12 klebsiella
26	7	3.1	195	16	Q9KCD7	Q9KCD7 bacillus ha
27	7	3.1	209	16	Q910U7	Q910U7 pseudomonas
28	7	3.1	209	16	Q9CJH4	Q9CJH4 pasteurella
29	7	3.1	211	16	Q67841	Q67841 aquifex aeo
30	7	3.1	220	2	Q9EMW7	Q9EMW7 streptomyces
31	7	3.1	225	16	Q97NE9	Q97NE9 streptococcus
32	7	3.1	230	16	Q9PHZ5	Q9PHZ5 campylobact
33	7	3.1	234	2	Q32330	Q32330 clostridium
34	7	3.1	240	12	Q9WH72	Q9WH72 thailand to
35	7	3.1	240	16	P73054	P73054 synechocyst
36	7	3.1	241	16	Q97KF4	Q97KF4 clostridium
37	7	3.1	247	16	Q9XDI6	Q9XDI6 neisseria m
38	7	3.1	247	16	Q9JX14	Q9JX14 neisseria m
39	7	3.1	251	16	Q9KOA1	Q9KOA1 neisseria m
40	7	3.1	255	16	Q92LR6	Q92LR6 rhizobium m
41	7	3.1	257	16	Q917C1	Q917C1 pseudomonas
42	7	3.1	259	2	Q9F728	Q9F728 pseudomonas
43	7	3.1	262	16	Q9PLI8	Q9PLI8 chlamydia m
44	7	3.1	270	5	Q9VDX1	Q9VDX1 drosophila
45	7	3.1	272	16	Q98PM7	Q98PM7 mycoplasma

## ALIGNMENTS

RESULT	ID	PRELIMINARY:	PRT:	364 AA.
1	Q9B0G4	Q9B0G4:		
AC	Q9B0G4:	Q9B0G4:		
DT	01-JUN-2001 (TREMBLrel. 17, Created)			
DT	01-JUN-2001 (TREMBLrel. 17, Last sequence update)			
DT	01-DEC-2001 (TREMBLrel. 19, Last annotation update)			
DE	HYPOTHEICAL 42.1 KDA PROTEIN (LYSOPHOSPHATIDIC ACID			
DE	ACYLTRANSFERASE-EPSILON) (EC 2.3.1.51).			
GN	DKEPZP61C222 OR LPAAT-E.			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
OX	NCBI_TaxID=9606;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=AMYGDALA;			
RX	MEDLINE=21154917; PubMed=11230166;			
RA	Wiemann S., Weil B., Wellenreuther R., Gassenhuber J., Glassl S.,			
RA	Ansorge W., Boecker M., Bloeker H., Bauersachs S., Blum H.,			
RA	Lauber J., Duesterhoeft A., Beyer A., Koehler K., Struck N.,			
RA	Mewes H.W., Oltjenwaelder B., Obermaier B., Tampe J., Heubner D.,			
RA	Wambutt R., Korn B., Klein M., Poustka A.;			
RT	"Towards a Catalog of Human Genes and Proteins: Sequencing and			
RT	Analysis of 500 Novel Complete Protein Coding Human cDNAs.";			
RT	Genome Res. 11:422-435(2001).			
[2]				
RA	SEQUENCE FROM N.A.			
RT	leung D.W.;			
RT	"Cloning and expression of LPAAT-epsilon.";			
RT	Submitted (MAY-2001) to the EMBL/Genbank/DBJ databases.			
DR	EMBL; AL136587; CAB65522.1; -			
DR	EMBL; AF375789; AAK54809.1; -			
DR	InterPro: IPR002123; Acyltransferase.			
DR	Pfam: PF01553; Acyltransferase; 1			
KW	Hypothetical protein; transferase; Acyltransferase.			
SO	SEQUENCE 364 AA: 42072 MW: 90A0F87FC78081 CRC64:			



KW Complete proteome.  
SQ SEQUENCE 264 AA; 30687 MW; DCE5727A07C74F7 CRC64;

Query Match 3.9%; Score 9; DB 16; Length 264;  
Best Local Similarity 100.0%; Pred. No. 0.47;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 157 LVIPPEGR 165  
|||||  
DB 156 LVIPPEGR 164

RESULT 5  
ID 095R12 PRELIMINARY; PRT; 386 AA.

DT 01-DEC-2001 (TREMBlrel. 19, Created)  
DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)  
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)  
DE HYPOTHEtical 44.0 KDA PROTEIN.

GN F28B3.9.  
OS Caenorhabditis elegans.  
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;  
OC Rhabditidae; Peloderinae; Caenorhabditis.  
OX NCBI\_Taxid=6239;

RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-BRISTOL N2;  
RX MEDLINE=99069613; PubMed=9851916;

RA None;  
RT "Genome sequence of the nematode C. elegans: a platform for investigating biology. The C. elegans Sequencing Consortium.";  
RL Science 282:2012-2018(1998).

RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN-BRISTOL N2;  
RA Geisel C., Kramer J., Smith A.;  
RT "The sequence of C. elegans cosmid F28B3.";  
RL Submitted (MAY-1997) to the EMBL/GenBank/DBJ databases.

RN [3]  
RP SEQUENCE FROM N.A.  
RC STRAIN-BRISTOL N2;  
RA Waterston R.;  
RT "Direct Submission.";  
RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.

DR EMBL; AF003136; AK93853.1; -;  
KW Hypothetical protein.  
SQ SEQUENCE 386 AA; 43978 MW; A1E3020C201452A6 CRC64;

Query Match 3.9%; Score 9; DB 5; Length 386;  
Best Local Similarity 100.0%; Pred. No. 0.66;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 211 LDATYDTV 219  
|||||  
DB 230 LDATYDTV 238

RESULT 6  
ID 098E87 PRELIMINARY; PRT; 115 AA.

DT 01-OCT-2001 (TREMBlrel. 18, Created)  
DT 01-OCT-2001 (TREMBlrel. 18, Last sequence update)  
DT 01-OCT-2001 (TREMBlrel. 18, Last annotation update)  
DE M14358 PROTEIN.

GN Rhizobium loti (Mesorhizobium loti).  
OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;  
OC Phyllobacteriaceae; Mesorhizobium.  
OX NCBI\_Taxid=381;

RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-MAFF303099;  
RX MEDLINE=21082930; PubMed=11214968;  
RA Kaneo T., Nakamura Y., Sato S., Asamizu E., Kato T., Sasamoto S.,  
RA Watanabe A., Idegawa K., Ishikawa A., Kawashima K., Kimura T.,  
RA Kishida Y., Kiyokawa C., Kohara M., Matsumoto M., Matsuno A.,  
RA Mochizuki Y., Nakayama S., Nakazaki N., Shimo S., Sugimoto M.,  
RA Takeuchi C., Yamada M., Tabata S.;  
RT "Complete genome structure of the nitrogen-fixing symbiotic bacterium Mesorhizobium loti.";  
RL DNA Res. 7:331-338(2000).  
DR EMBL; AP003004; BAB51033.1; -;  
KW Complete proteome.  
SQ SEQUENCE 115 AA; 12335 MW; 6AF43D6451F42BE8 CRC64;

Query Match 3.5%; Score 8; DB 16; Length 115;  
Best Local Similarity 100.0%; Pred. No. 2.6;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 174 LSASOAF 181  
|||||  
DB 31 LSASOAF 38

RESULT 7  
ID 09A675 PRELIMINARY; PRT; 236 AA.

DT 01-JUN-2001 (TREMBlrel. 17, Created)  
DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)  
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)  
DE 1-ACYL-SN-GLYCEROL-3-PHOSPHATE ACYLTRANSFERASE.

GN CC2219.  
OS Caulobacter crescentus.  
OC Bacteria; Proteobacteria; alpha subdivision; Caulobacter group;  
OC Caulobacter.  
OX NCBI\_Taxid=69394;

RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-ATCC 19089 / CB15;  
RX MEDLINE=21173698; PubMed=11259647;  
RA Nierman W.C., Feldblyum T.V., Laub M.T., Paulsen I.T., Nelson K.E.,  
RA Eisen J., Heidelberg J.F., Alley M.R.K., Ohta N., Maddock J.R.,  
RA Potocka I., Nelson W.C., Newton A., Stephens C., Phadke N.D., Ely B.,  
RA Kolonay J.F., Smit J., Craven M.B., Khouri H., Shetty J., Berry K.,  
RA Uterback T., Tran K., Wolf A., Vamathevan J., Ermolaeva M., White O.,  
RA Salzberg S.L., Venter J.C., Shapiro L., Fraser C.M.;  
RT "Complete genome sequence of Caulobacter crescentus.";  
RL Proc. Natl. Acad. Sci. U.S.A. 98:4136-4141(2001).

DR TIGR; CC2219; -;  
DR InterPro: IPR002123; Acyltransferase.  
DR Pfam: PF01553; Acyltransferase; 1.  
KW Transferase; Acyltransferase; Complete proteome.  
SQ SEQUENCE 236 AA; 26362 MW; 83694D3966BCA1A0 CRC64;

Query Match 3.5%; Score 8; DB 16; Length 236;  
Best Local Similarity 100.0%; Pred. No. 4.9;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 158 VIFPEGR 165  
|||||  
DB 146 VIFPEGR 153

RESULT 8  
ID 084743 PRELIMINARY; PRT; 262 AA.

DT 01-OCT-2001 (TREMBlrel. 18, Created)  
DT 01-OCT-2001 (TREMBlrel. 18, Last sequence update)  
DT 01-OCT-2001 (TREMBlrel. 18, Last annotation update)  
DE M14358 PROTEIN.

GN Rhizobium loti (Mesorhizobium loti).  
OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;  
OC Phyllobacteriaceae; Mesorhizobium.  
OX NCBI\_Taxid=381;

DT 01-NOV-1998 (TREMBLrel. 08, Created)  
 DT 01-NOV-1998 (TREMBLrel. 08, Last sequence update)  
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)  
 DE METAL DEPENDENT HYDROLASI.  
 GN YJCJ OR CT738.  
 OS Chlamydia trachomatis.  
 OC Bacteria; Chlamydiales; Chlamydiaceae; Chlamydia.  
 OX NCBI\_TaxID=813;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=D/UV-3/CX;  
 RX MEDLINE=99000609; PubMed=9784136;  
 RA Stephens R.S., Kaiman S., Lammell C.J., Fan J., Marathe R., Aravind L.,  
 RA Mitchell W.P., Olinger L., Tatusov R.L., Zhao Q., Koonin E.V.,  
 RA Davis R.W.;  
 RT "Genome sequence of an obligate intracellular pathogen of humans:  
 RT Chlamydia trachomatis."  
 RL Science 282:754-759(1998).  
 DR EMBL: AE01344; AAC68333.1; -.  
 KW Hydrolase; Complete proteome.  
 SO SEQUENCE 262 AA; 29082 MW; E36A6298A23CEBF CRC64;

Query Match 3.5%; Score 8; DB 16; Length 262;  
 Best Local Similarity 100.0%; Pred. No. 5.4;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 191 HVLPRIK 198  
 DB 202 HVLPRIK 209

RESULT 9  
 ID 083380 PRELIMINARY; PRT; 282 AA.  
 AC 083380:  
 DT 01-NOV-1998 (TREMBLrel. 08, Created)  
 DT 01-NOV-1998 (TREMBLrel. 08, Last sequence update)  
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)  
 DE LYSOPHOSPHATIDIC ACID ACYLTRANSFERASE, PUTATIVE.  
 GN TP0361.  
 OS Treponema pallidum.  
 OC Bacteria; Spirochaetales; Spirochaetaceae; Treponema.  
 OX NCBI\_TaxID=160;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=NICHOLS;  
 RX MEDLINE=98332770; PubMed=9665876;  
 RA Fraser C.M., Norris S.J., Weinstock G.M., White O., Sutton G.G.,  
 RA Dodson R., Gwinn M., Hickey E.K., Clayton R., Ketchum K.A.,  
 RA Sodergren E., Hardham J.M., McLeod M.P., Salzberg S., Peterson J.,  
 RA Khalak H., Richardson D., Howell J.K., Chidambaram M., Uterback T.,  
 RA McDonald L., Artlich P., Bowman C., Cotton M.D., Fujii C., Garland S.,  
 RA Hatch B., Horst K., Roberts K., Sandusky M., Weidman J., Smith H.O.,  
 RA Venter J.C.;  
 RT "Complete genome sequence of Treponema pallidum, the syphilis  
 RT spirochete."  
 RL Science 281:375-388(1998).  
 DR EMBL: AE001215; AAC65346.1; -.  
 DR TIGR: TP0361; -.  
 DR InterPro: IPR002123; Acyltransferase.  
 DR Pfam: PF01553; Acyltransferase; 1.  
 KW Transferase; Acyltransferase; Complete proteome.  
 SO SEQUENCE 282 AA; 31700 MW; DEF69003CD6C61AD CRC64;

Query Match 3.5%; Score 8; DB 16; Length 282;  
 Best Local Similarity 100.0%; Pred. No. 5.8;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 158 VIFPESTR 165  
 DB 170 VIFPESTR 177

RESULT 10  
 ID 09P5W2 PRELIMINARY; PRT; 454 AA.  
 AC 09P5W2:  
 DT 01-OCT-2000 (TREMBLrel. 15, Created)  
 DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)  
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)  
 DE RELATED TO GUANINE DEAMINASE.  
 GN B3E4.200.  
 OS Neurospora crassa.  
 OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;  
 OC Sordariales; Sordariaceae; Neurospora.  
 OX NCBI\_TaxID=5141;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Schulte U., Algn V., Hohelsel J., Brandt P., Fartmann B., Holland R.,  
 RA Nyakatura G., Mewes H.W., Mannhaupt G.;  
 RL Submitted (MAY-2000) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RA German Neurospora genome project;  
 RL Submitted (NOV-2001) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: AL355931; CAB91408.2; -.  
 SO SEQUENCE 454 AA; 49778 MW; 8C1EC14BAECA421D CRC64;

Query Match 3.5%; Score 8; DB 3; Length 454;  
 Best Local Similarity 100.0%; Pred. No. 8.9;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 189 LKHVTPR 196  
 DB 208 LKHVTPR 215

RESULT 11  
 ID 09KF10 PRELIMINARY; PRT; 458 AA.  
 AC 09KF10:  
 DT 01-OCT-2000 (TREMBLrel. 15, Created)  
 DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)  
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)  
 DE RNA METHYLTRANSFERASE.  
 GN BHO687.  
 OS Bacillus halodurans.  
 OC Bacteria; Firmicutes; Bacillus/Clostridium group;  
 OC Bacillus/Staphylococcus group; Bacillus.  
 OX NCBI\_TaxID=86655;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=C-125 / JCM 9153;  
 RX MEDLINE=20512582; PubMed=11058132;  
 RA Takami H., Nakasone K., Takaki Y., Maeno G., Sasaki R., Masui N.,  
 RA Fuji F., Hirama C., Nakamura Y., Ogasawara N., Kuhara S.,  
 RA Horikoshi K.;  
 RT "Complete genome sequence of the alkaliphilic bacterium Bacillus  
 RT halodurans and genomic sequence comparison with Bacillus subtilis."  
 RL Nucleic Acids Res. 28:4317-4331(2000).  
 DR EMBL: AF001509; BAB04406.1; -.  
 DR InterPro: IPR000051; SAM\_bind.  
 DR InterPro: IPR001566; TRNA\_1.  
 DR PROSITE: PS01230; TRNA\_1; UNKNOWN\_1.  
 KW Transferase; Methyltransferase; Complete proteome.  
 SO SEQUENCE 458 AA; 51555 MW; 3D3A661B8891240 CRC64;

Query Match 3.5%; Score 8; DB 16; Length 458;  
 Best Local Similarity 100.0%; Pred. No. 8.9;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 167 NPEQTKVL 174



Db 291 NPEQTKVL 298

```

|||||
RESULT 12
ID 031503 PRELIMINARY; PRT; 459 AA.
AC 031503;
DT 01-JAN-1998 (TREMBLrel. 05, Created)
DT 01-JAN-1998 (TREMBLrel. 05, last sequence update)
DE 01-OCT-2001 (TREMBLrel. 18, last annotation update)
DE YEPA PROTEIN.
GN YEPA.
OS Bacillus subtilis.
OC Bacteria; Firmicutes; Bacillus/Clostridium group;
OC Bacillus/Staphylococcus group; Bacillus.
OX NCBI_TaxID=1423;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=168;
RX MEDLINE=98044033; PubMed=9384377;
RA Kunst F., Ogasawara N., Moszer I., Albertini A.M., Alloni G.,
RA Azevedo V., Bertero M.G., Bessieres P., Bolochin A., Borchert S.,
RA Borris R., Boursier L., Brans A., Braun M., Brignell S.C., Bron S.,
RA Brouillet S., Bruschi C.V., Caldwell B., Capuano V., Carter N.M.,
RA Choi S.K., Codani J.J., Connerion I.F., Cummings N.J., Daniel R.A.,
RA Denizot F., Devine K.M., Dusterhoft A., Ehrlich S.D., Emerson P.T.,
RA Erlan K.D., Errington J., Fabret C., Ferrari E., Foulger D.,
RA Fritz C., Fujita M., Fujita Y., Fuma S., Galizzi A., Galleron N.,
RA Ghim S.Y., Glaser P., Goffeau A., Gollightly E.J., Grandi G.,
RA Giuseppe G., Guy B.J., Haga K., Halech J., Harwood C.R., Henaut A.,
RA Hilbert H., Holsappel S., Hosono S., Hullo M.F., Itaya M., Jones L.,
RA Joris B., Karamata D., Kasahara Y., Klaerr-Blanchard M., Klein C.,
RA Kobayashi Y., Koetter P., Konigstein G., Krogh S., Kumano M.,
RA Kurita K., Lapidus A., Lardinois S., Lauber J., Lazarevic V.,
RA Lee S.M., Levine A., Liu H., Masuda S., Mauel C., Medigue C.,
RA Medina N., Meliado R.P., Mizuno M., Moesti D., Nakai S., Noback M.,
RA Noone D., O'Reilly M., Ogawa K., Ogilwara A., Oudega B., Park S.H.,
RA Paro V., Pohl T.M., Portelle D., Porwollik S., Prescott A.M.,
RA Prescan E., Pujic P., Purnelle B., Rapoport G., Rey M., Reynolds S.,
RA Rieger M., Rivolta C., Rocha E., Roche B., Rose M., Sadate I.,
RA Sato T., Scanlan E., Schleich S., Schroeter R., Scoffone F.,
RA Sekiguchi J., Sekowska A., Seror S.J., Serro P., Shit B.S., Soldo B.,
RA Sorokin A., Tacconi E., Takagi T., Takahashi H., Takemaru K.,
RA Takeuchi M., Tanakoshi A., Tanaka T., Terpstra P., Tognoni A.,
RA Tosato V., Uchiyama S., Vandenbol M., Vannier F., Vassariotti A.,
RA Viari A., Wanduit R., Wedler E., Wedler H., Wellzenegger T.,
RA Winters P., Wipat A., Yamamoto H., Yamane K., Yasumoto K., Yata K.,
RA Yoshida K., Yoshikawa H.F., Zumstein E., Yoshikawa H., Danchin A.;
RT "The complete genome sequence of the gram-positive bacterium Bacillus
RT subtilis."
RL Nature 390:249-256(1997).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=168;
RA Kunst F., Ogasawara N., Yoshikawa H., Danchin A.;
RL Submitted (NOV-1997) to the EMBL/Genbank/DBJ databases.
DR EMBL: Z99107; CAB12493.1; -.
DR InterPro: IPR002792; DUF90.
DR InterPro: IPR000051; SAM_bind.
DR InterPro: IPR001566; TRMA_1.
DR Pfam: PF01938; TRAM_1.
DR PROSITE: PS01230; TRMA_1; UNKNOWN_1.
DR PROSITE: PS01231; TRMA_2; 1.
KW Complete proteome.
SQ SEQUENCE 459 AA; 51819 MW; 35E9954779421D5 CRC64;

```

Query Match 3.5%; Score 8; DB 16; Length 459;  
 Best Local Similarity 100.0%; Pred. No. 9;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 167 NPEQTKVL 174

Db 292 NPEQTKVL 299

```

|||||
RESULT 13
ID 049927 PRELIMINARY; PRT; 483 AA.
AC 049927;
DT 01-JUN-1998 (TREMBLrel. 06, Created)
DT 01-JUN-1998 (TREMBLrel. 06, last sequence update)
DE 01-DEC-2001 (TREMBLrel. 19, last annotation update)
DE P54 PROTEIN.
GN P54.
OS Pisum sativum (Garden pea).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids I; Fabales; Fabaceae; Papilionoideae; Viciae; Pisum.
OX NCBI_TaxID=3888;
RN [1]
RP SEQUENCE FROM N.A.
RA Castillo J., Marquez J.A., Franco L., Ballestar E., Rodrigo M.I.;
RL Submitted (FEB-1997) to the EMBL/Genbank/DBJ databases.
DR EMBL: Y11207; CAA72090.1; -.
DR HSP: P50477; ICAU.
DR InterPro: IPR001113; Seedstore_7s.
DR Pfam: PF00546; Seedstore_7s; 1.
DR Pfam: PF02808; Seedstore_7s.C; 1.
SQ SEQUENCE 483 AA; 54662 MW; 8127BDAA0178F3D CRC64;

```

Query Match 3.5%; Score 8; DB 10; Length 483;  
 Best Local Similarity 100.0%; Pred. No. 9.4;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 151 ACTPMYLV 158  
 Db 184 ACTPMYLV 191

```

RESULT 14
ID Q9VVG0 PRELIMINARY; PRT; 1820 AA.
AC Q9VVG0;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, last sequence update)
DT 01-JUN-2001 (TREMBLrel. 17, last annotation update)
DE PUTATIVE CADHERIN PRECURSOR (CG6445 PROTEIN).
GN CG6445.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BERKELEY;
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celinker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galie R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe C.R., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson S.M., Miklos G.L.G.,
RA Abill J.F., Agdayanl A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.A., Bouck J., Brokstein P., Brotlier P.,
RA Burris K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablo B., Delcher A., Deng Z., Deyes A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferritara S., Fleischmann W.,
RA Foster C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,

```

RA Glodok A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,  
RA Harris N.L., Harvey D., Helman T.J., Hernandez J.R., Houck J.,  
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,  
RA Jalali M., Kalush F., Karpén G.H., Ke Z., Kennison J.A., Ketchum K.A.,  
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,  
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,  
RA Liu X., Matel B., McIntosh T.C., McLeod M.P., McPherson D.,  
RA Merkulov G., Mishina N.V., Modyarty C., Morris J., Moshrefi A.,  
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,  
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,  
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,  
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,  
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,  
RA Splier E., Spradling A.C., Stapleton M., Strong R., Sun E.,  
RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,  
RA Wang Z.-Y., Maassman D.A., Weinstein G.M., Weissbach J.,  
RA Williams S.M., Woodage T., Morley K.C., Wu D., Yang S., Yao Q.A.,  
RA Ye J., Yeh R.F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,  
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,  
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;  
RT "The genome sequence of *Drosophila melanogaster*.";  
RL Science 287:2185-2195(2000).  
CC -1- SIMILARITY: BELONGS TO THE CADHERIN FAMILY.  
CC -1- FUNCTION: MAY BE INVOLVED IN CELL ADHESION.  
DR EMBL: AE003524; AAF49351.1; -.  
DR HSSP: P15116; INCU.  
DR FLYBase: FBgn0036715; CG6445.  
DR InterPro: IPR002126; Cndherin.  
DR Pfam: PF00028; cndherin.12.  
DR PRINTS: PRO0205; CADHERIN.  
DR SMART: SM00112; CA: 13.  
DR PROSITE: PS00232; CADHERIN\_1; 1.  
DR PROSITE: PS50268; CADHERIN\_2; 13.  
KW Hypothetical protein; Cell adhesion; Glycoprotein; Repeat; Signal;  
KW Calcium-binding.  
FT SIGNAL 1 14  
FT CHAIN 15 1820 POTENTIAL.  
FT DOMAIN 36 128 PUTATIVE CADHERIN.  
FT DOMAIN 137 245 CADHERIN 1.  
FT DOMAIN 254 365 CADHERIN 2.  
FT DOMAIN 374 483 CADHERIN 3.  
FT DOMAIN 492 585 CADHERIN 4.  
FT DOMAIN 590 696 CADHERIN 5.  
FT DOMAIN 718 815 CADHERIN 6.  
FT DOMAIN 824 939 CADHERIN 7.  
FT DOMAIN 955 1053 CADHERIN 8.  
FT DOMAIN 1062 1170 CADHERIN 9.  
FT DOMAIN 1170 1284 CADHERIN 10.  
FT DOMAIN 1295 1394 CADHERIN 11.  
FT DOMAIN 1407 1517 CADHERIN 12.  
FT DOMAIN 1517 191 CADHERIN 13.  
FT CARBOHYD 191 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 371 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 435 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 678 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 727 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 868 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 890 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 1040 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 1111 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 1112 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 1233 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 1294 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 1419 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 1452 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 1557 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 1584 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 1610 N-LINKED (GLCNAC. . .) (POTENTIAL).  
SQ SEQUENCE 1820 AA; 200996 MW; 8D61B907FA15C155 CRC64;

Query Match 3.5%; Score 8; DB 5; Length 1820;  
Best Local Similarity 100.0%; Pred. No. 31;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 102 GHVRYVLK 109  
Db 516 GHVRYVLK 523

RESULT 15  
ID 028219 PRELIMINARY; PRT; 54 AA.  
AC 028219;  
DT 01-NOV-1996 (TREMblrel. 01, Created)  
DT 01-DEC-2001 (TREMblrel. 19, Last sequence update)  
DE HEPARIN BINDING EGF-LIKE GENE PRECURSOR.  
OS Cercopithecus aethiops (Green monkey) (Griwet).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;  
OC Cercopithecinae; Cercopithecus.  
OX NCBI\_Taxid=9534;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=KIDNEY;  
RX MEDLINE=97444292; PubMed=9300824;  
RT Loukianov E.V., Loukianov T.I., Mledlocha A., Olunes S.;  
RT "Expression of mRNA for a short form of heparin-binding EGF-like  
RT growth factor.";  
RL Gene 195:81-86(1997).  
DR EMBL: X89729; CAA61861.1; -.  
KW Signal.  
FT SIGNAL 1 23 POTENTIAL.  
FT SEQUENCE 54 AA; 5888 MW; 7C0E67E85695873F CRC64;

Query Match 3.1%; Score 7; DB 6; Length 54;  
Best Local Similarity 100.0%; Pred. No. 15;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 4 LPSVVL 10  
Db 3 LPSVVL 9

Search completed: August 28, 2002, 11:28:35  
Job time: 548 sec

---

THIS PAGE BLANK (088470)

6

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 28, 2002, 11:20:09 ; Search time 39.16 Seconds  
(without alignments)  
559.458 Million cell updates/sec

Title: US-09-853-526-70

Perfect score: 228  
Sequence: 1 MRVLPSTVVLGTAPTYVLA.....NYLDIAIVDVVYEGKDDGG 228

Scoring table:  
Gapop 60.0 , Gapext 60.0

Searched: 283138 seqs, 96089334 residues

Word size : 0

Total number of hits satisfying chosen parameters: 283138

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database :  
1: PIR1:\*  
2: PIR2:\*  
3: PIR3:\*  
4: PIR4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	9	3.9	237	2	B71827
2	9	3.9	240	2	D64688
3	9	3.9	264	2	E82903
4	9	3.9	918	2	T34057
5	8	3.5	236	2	H90892
6	8	3.5	236	2	B87524
7	8	3.5	239	2	A85725
8	8	3.5	262	2	A71478
9	8	3.5	282	2	G71334
10	8	3.5	396	2	S54641
11	8	3.5	424	2	T49607
12	8	3.5	458	2	G83735
13	8	3.5	459	2	E69793
14	8	3.5	483	2	T06459
15	8	3.5	611	2	A85000
16	7	3.1	54	2	S58122
17	7	3.1	57	2	D86108
18	7	3.1	117	2	I81209
19	7	3.1	126	2	I81210
20	7	3.1	141	2	S58117
21	7	3.1	143	2	B90532
22	7	3.1	148	2	C86732
23	7	3.1	194	2	PC1136
24	7	3.1	195	2	C83854
25	7	3.1	208	1	A38432
26	7	3.1	208	1	A41914
27	7	3.1	208	1	UC1409
28	7	3.1	209	2	B83329
29	7	3.1	210	2	A10342

30	7	3.1	211	2	E70476	2-acylglycerophosph
31	7	3.1	225	2	H95244	conserved hypothet
32	7	3.1	226	2	E98109	conserved hypothet
33	7	3.1	230	2	E81397	probable 1-acylgly
34	7	3.1	240	2	S75162	hypothetical prote
35	7	3.1	241	2	B97019	1-acyl-sn-glycerol
36	7	3.1	247	2	A81957	1-acylglycerol-3-p
37	7	3.1	247	2	G81013	1-acyl-sn-glycerol
38	7	3.1	250	2	T27772	26S proteasome reg
39	7	3.1	250	2	E70104	1-acylglycerol-3-p
40	7	3.1	250	2	AF3384	outer membrane pro
41	7	3.1	251	2	A81167	conserved hypothet
42	7	3.1	257	2	A83645	probable acyltrans
43	7	3.1	262	2	H81740	conserved hypothet
44	7	3.1	268	2	AH2521	hypothetical prote
45	7	3.1	269	2	G91169	probable acyltrans

ALIGNMENTS

RESULT 1  
B71827  
Probable 1-acylglycerol-3-phosphate O-acyltransferase (EC 2.3.1.51) - Helicobacter pylori  
C:Species: Helicobacter pylori  
A:Variate: strain J99  
C:Date: 12-Feb-1999 #sequence-revision 12-Feb-1999 #text-change 06-Oct-2000  
C:Accession: B71827  
R:Alm, R.A.; Ling, L.S.L.; Molr, D.T.; King, B.L.; Brown, E.D.; Doig, P.C.; Smith, D.; Ives, C.; Gibson, R.; Merberg, D.; Mills, S.D.; Jlang, Q.; Taylor, D.E.; Vovis, G.F. Nature 397, 176-180, 1999  
A:Title: Genomic sequence comparison of two unrelated isolates of the human gastric P  
A:Reference number: A71800; MUID:99120557  
A:Accession: B71827  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-237 <ARN>  
A:Cross-references: GB:AE001550; GB:AE001439; NID:g4155872; PIDN:AAD06852.1; PID:g415  
A:Experimental source: strain J99  
C:Genetics:  
A:Gene: plsc  
C:Superfamily: mouse 1-acylglycerol-3-phosphate O-acyltransferase  
C:Keywords: acyltransferase; coenzyme A

Query Match 3.9%; Score 9; DB 2; Length 237;  
Best Local Similarity 100.0%; Pred. No. 0.27;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 157 LVTFPEGTR 165  
DB 142 LVTFPEGTR 150

RESULT 2  
D64688  
Probable 1-acylglycerol-3-phosphate O-acyltransferase (EC 2.3.1.51) - Helicobacter pylori  
C:Species: Helicobacter pylori  
C:Date: 09-Aug-1997 #sequence-revision 09-Aug-1997 #text-change 06-Oct-2000  
C:Accession: D64688  
R:Tomp, J.F.; White, O.; Kerlavage, A.R.; Clayton, R.A.; Sutton, G.G.; Fleischmann, R.  
Peterson, S.; Loftus, B.; Richardson, D.; Dodson, R.; Khaliq, H.G.; Glodek, A.; McK  
son, J.D.; Kelley, J.M.; Cotton, M.D.; Weidman, J.M.; Fujii, C.; Bowman, C.; Watney,  
Nature 388, 539-547, 1997  
A:Authors: Wallin, E.; Hayes, W.S.; Borodovsky, M.; Karpk, P.D.; Smith, H.O.; Fraser,  
A:Title: The complete genome sequence of the gastric pathogen Helicobacter pylori.  
A:Reference number: A64520; MUID:97394467  
A:Accession: D64688  
A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-240 <TOM>  
A:Cross-references: GB:AE000636; GB:AE000511; NID:g214517; PIDN:AAD08393.1; PID:g231  
C:Superfamily: mouse 1-acylglycerol-3-phosphate O-acyltransferase

C:Keywords: acyltransferase; coenzyme A

Query Match 3.9%; Score 9; DB 2; Length 240;  
Best Local Similarity 100.0%; Pred. No. 0.27;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 157 LVIRPECTR 165  
|||||  
DB 142 LVIRPECTR 150

RESULT 3  
E82903  
1-acyl-sn-glycerol-3-phosphate acyltransferase U0344 [imported] .. Ureaplasma urealyticum  
C:Species: Ureaplasma urealyticum  
C>Date: 18-Aug-2000 #sequence\_revision 20-Aug-2000 #text\_change 20-Aug-2000  
C:Accession: E82903  
R:Glass, J.I.; Lefkowitz, E.J.; Glass, J.S.; Helner, C.R.; Chen, E.Y.; Cassell, G.H.  
submitted to GenBank, February 2000  
A:Description: The complete sequence of Ureaplasma urealyticum: Alternate views of a mlr  
A:Reference number: A82870  
A:Accession: E82903  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-264 <GUA>  
A:Cross-references: GB:AE002131; GB:AF222894; NID:96899316; PIDN:AAF30753.1; GSPDB:GN001  
A:Experimental source: serovar 3; biovar 1  
C:Genetics:  
A:Gene: plsc; U0344  
A:Genetic code: SGC3

Query Match 3.9%; Score 9; DB 2; Length 264;  
Best Local Similarity 100.0%; Pred. No. 0.3;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 157 LVIRPECTR 165  
|||||  
DB 156 LVIRPECTR 164

RESULT 4  
T34057  
hypothetical protein F28B3.5 - Caenorhabditis elegans  
C:Species: Caenorhabditis elegans  
C>Date: 29-Oct-1999 #sequence\_revision 29-Oct-1999 #text\_change 29-Oct-1999  
C:Accession: T34057  
R:Geisler, C.; Kramer, J.; Smith, A.  
submitted to the EMBL Data Library, May 1997  
A:Description: The sequence of C. elegans cosmid F28B3.  
A:Reference number: 221469  
A:Accession: T34057  
A:Status: preliminary; translated from GB/EMBL/DBST  
A:Molecule type: DNA  
A:Residues: 1-918 <GEI>  
A:Cross-references: EMBL:AF003136; PIDN:AA893636.1; GSPDB:GN00019; CESP:F28B3.5  
A:Experimental source: strain Bristol N2; clone F28B3  
C:Genetics:  
A:Gene: CESP:F28B3.5  
A:Map position: 1  
A:Intons: 85/3; 129/3; 235/3; 418/2; 482/3; 532/3; 736/3; 829/3; 850/2

Query Match 3.9%; Score 9; DB 2; Length 918;  
Best Local Similarity 100.0%; Pred. No. 0.9;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 211 LDAIVDVT 219  
|||||  
DB 762 LDAIVDVT 770

RESULT 5  
H90892  
probable fimbrial chaperone protein precursor [imported] - Escherichia coli (strain O  
C:Species: Escherichia coli  
C>Date: 18-Jul-2001 #sequence\_revision 18-Jul-2001 #text\_change 03-Aug-2001  
C:Accession: H90892  
R:Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.  
gasawara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shingawa, H.  
DNA Res. 8, 11-22, 2001  
A:Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and g  
A:Reference number: A99629; M01D:21156231; PMID:11258796  
A:Accession: H90892  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-236 <NAV>  
A:Cross-references: GB:BA000007; PIDN:BA835535.1; PID:q13361578; GSPDB:GN00154  
A:Experimental source: strain O157:H7, substrain RIMD 0509952  
C:Genetics:  
A:Gene: ECG2112  
C:Superfamily: chaperone protein papd

Query Match 3.5%; Score 8; DB 2; Length 234;  
Best Local Similarity 100.0%; Pred. No. 3;  
Matches 8; Conservative 0; Mismatches 0; Indels 0;

OY 193 LTPRIKAT 200  
|||||  
DB 227 LTPRIKAT 234

RESULT 6  
B87524  
1-acyl-sn-glycerol-3-phosphate acyltransferase [imported] - Caulobacter crescentus  
C:Species: Caulobacter crescentus  
C>Date: 20-Apr-2001 #sequence\_revision 20-Apr-2001 #text\_change 20-Apr-2001  
C:Accession: B87524  
R:Nierman, W.C.; Feldbylum, T.V.; Paulsen, I.T.; Nelson, K.E.; Eisen, J.; Heidelberg,  
B.; Laub, M.T.; Deboy, R.T.; Dodson, R.J.; Durkin, A.S.; Gwinn, M.L.; Haft, D.H.; Ko  
n, J.; Ermolaeva, M.; White, O.; Salzberg, S.L.; Shapiro, L.; Venter, J.C.; Fraser, C  
Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001  
A:Title: Complete Genome Sequence of Caulobacter crescentus.  
A:Reference number: A87249; M01D:21173698; PMID:11259647  
A:Accession: B87524  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-236 <STO>  
A:Cross-references: GB:AE005673; NID:q13423724; PIDN:AAK24190.1; GSPDB:GN00148  
C:Genetics:  
A:Gene: CC2219

Query Match 3.5%; Score 8; DB 2; Length 236;  
Best Local Similarity 100.0%; Pred. No. 3;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 158 VIRPECTR 165  
|||||  
DB 146 VIRPECTR 153

RESULT 7  
A85725  
probable fimbrial chaperone protein 22201 [imported] - Escherichia coli (strain O157:  
C:Species: Escherichia coli  
C>Date: 16-Feb-2001 #sequence\_revision 16-Feb-2001 #text\_change 14-Sep-2001  
C:Accession: A85725  
R:Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; May  
hiller, L.; Grobeck, E.J.; Davis, N.W.; Llin, A.; Dimalanta, E.; Potamousis, K.; Apoda  
Nature 409, 529-533, 2001  
A:Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.  
A:Reference number: A85480; M01D:21074935; PMID:11206551  
A:Accession: A85725

THIS PAGE IS BLANK (USP)

A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-239 <STO>  
A:Cross-references: GB:AE005174; NID:g12515164; PIDN:AAG56261.1; GSPDB:GN00145; UMGF:222  
A:Experimental source: strain 0157:H7, substrain EDJ933  
C:Genetics:  
A:Gene: 22201  
C:Superfamily: chaperone protein papp

Query Match 3.5%; Score 8; DB 2; Length 239;  
Best Local Similarity 100.0%; Pred. No. 3;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 193 LTPRIKAT 200  
DB 230 LTPRIKAT 237

#### RESULT 8

A71478

probable metal dependent hydrolase - Chlamydia trachomatis (serotype D, strain UW3/Cx)

C:Species: Chlamydia trachomatis

C:Date: 13-Sep-1998 #sequence\_revision 13-Sep-1998 #text\_change 08-Oct-1999

C:Accession: A71478

R:Stephens, R.S.; Kalman, S.; Lammel, C.J.; Fan, J.; Marathe, R.; Aravind, L.; Mitchell,

Science 282, 754-759, 1998

A:Title: Genome sequence of an obligate intracellular pathogen of humans: Chlamydia trach

A:Reference number: A71570; MUID:99000809

A:Accession: A71478

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-262 <ARN>

A:Cross-references: GB:AE001344; GB:AE001273; NID:g3329188; PIDN:AAC68333.1; PID:g332919

A:Experimental source: serotype D, strain UW-3/Cx

C:Genetics:

A:Gene: yycJ

Query Match 3.5%; Score 8; DB 2; Length 262;  
Best Local Similarity 100.0%; Pred. No. 3.3;

Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 191 HVLTPRIK 198

DB 202 HVLTPRIK 209

#### RESULT 9

G71334

probable lysophosphatidic acid acyltransferase - syphilis spirochete

C:Species: Treponema pallidum subsp. pallidum (syphilis spirochete)

C:Date: 24-Jul-1998 #sequence\_revision 24-Jul-1998 #text\_change 05-Nov-1999

C:Accession: G71334

R:Fraser, C.M.; Norris, S.J.; Weinstein, G.M.; White, O.; Sutton, G.G.; Dodson, R.; Gwin

ton, J.; Khailak, H.; Richardson, D.; Howell, J.R.; Chidambaram, M.; Uitterback, T.; McD

they, L.; Weidman, J.; Smith, H.O.; Venter, J.C.

A:Title: Complete genome sequence of Treponema pallidum, the syphilis spirochete.

A:Reference number: A71250; MUID:9832770

A:Accession: G71334

A:Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-282 <COT>

A:Cross-references: GB:AE001215; GB:AE000520; NID:g3322631; PIDN:AAG55346.1; PID:g332263

A:Experimental source: strain Nichols

C:Genetics:

A:Gene: TP0361

OY 158 VIFPECTR 165  
DB 170 VIFPECTR 177

#### RESULT 10

S54641

probable membrane protein YDR018c - yeast (Saccharomyces cerevisiae)

N:Alternate names: hypothetical protein D3246; hypothetical protein PZF396; hypotheti

C:Species: Saccharomyces cerevisiae

C:Date: 08-Jul-1995 #sequence\_revision 01-Sep-1995 #text\_change 20-Jun-2000

C:Accession: S54641; S63425; S67831; S72116

R:Dedman, K.; Brown, D.; Hamlyn, N.; Bowman, S.

submitted to the EMBL Data Library, May 1995

A:Reference number: S54638

A:Accession: S54641

A:Molecule type: DNA

A:Residues: 1-396 <DED>

A:Cross-references: EMBL:249770; NID:g840867; PIDN:CAA9843.1; PID:g840871

A:Experimental source: strain AB972

R:Eide, L.G.; Sander, C.; Prydz, H.

submitted to the EMBL Data Library, February 1996

A:Description: Sequencing and analysis of a 35.4 kb region on the left arm of chromos

A:Reference number: S63416

A:Accession: S63425

A:Molecule type: DNA

A:Residues: 1-396 <EID>

A:Cross-references: EMBL:X95966; NID:g1216215; PIDN:CAA65210.1; PID:g1216225

R:Prydz, H.; Eide, L.G.

submitted to the Protein Sequence Database, July 1996

A:Reference number: S67822

A:Accession: S67831

A:Molecule type: DNA

A:Residues: 1-396 <PRX>

A:Cross-references: EMBL:274314; NID:g1431443; PIDN:CAA98838.1; PID:g1431444; MIPS:YD

A:Experimental source: strain S288C

R:Eide, L.G.; Sander, C.; Prydz, H.

Yeast 12, 1085-1090, 1996

A:Title: Sequencing and analysis of a 35.4 kb region on the left arm of chromosome IV

A:Reference number: S72107; MUID:97051598

A:Accession: S72116

A:Status: nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-396 <EIV>

A:Cross-references: EMBL:X95966; NID:g1216215; PIDN:CAA65210.1; PID:g1216225

A:Note: the nucleotide sequence was submitted to the EMBL Data Library, February 1996

C:Genetics:

A:Map position: 4R

A:Note: YDR018c

C:Superfamily: probable membrane protein YBR042c

C:Keywords: transmembrane protein

F:27-43/Domain: transmembrane #status predicted <TM1>

F:69-85/Domain: transmembrane #status predicted <TM2>

F:376-392/Domain: transmembrane #status predicted <TM3>

Query Match 3.5%; Score 8; DB 2; Length 396;  
Best Local Similarity 100.0%; Pred. No. 4.7;

Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 211 LDAIYDVT 218

DB 265 LDAIYDVT 272

#### RESULT 11

T49607

guanine deaminase (Gde) related protein [Imported] - Neurospora crassa

N:Alternate names: protein B3E4.200

C:Species: Neurospora crassa

C:Date: 02-Jun-2000 #sequence\_revision 02-Jun-2000 #text\_change 02-Jun-2000

C:Accession: T49607

R;Schulte, U.; Aljun, V.; Hehnel, J.; Brandt, P.; Fartmann, B.; Holland, R.; Nyakatura, submitted to the Protein Sequence Database, May 2000

A:Reference number: 225022

A:Accession: T49607

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-424 <SCH>

A:Cross-references: EMBL:AJ355931; GSPDB:GN00116; NCSP:B3E4.200

A:Experimental source: BAC clone B3E4; strain OR74A

C:Genetics:

A:Gene: NCSP:B3E4.200

A:Map position: 6

Query Match 3.5%; Score 8; DB 2; Length 424;  
Best Local Similarity 100.0%; Pred. No. 5;

Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 189 LKHVLTFR 196  
|||||

DB 178 LKHVLTFR 185

#### RESULT 12

RNA methyltransferase BH0687 [imported] - *Bacillus halodurans* (strain C-125)

C:Species: *Bacillus halodurans*

C>Date: 01-Dec-2000 #sequence\_revision 01-Dec-2000 #text\_change 15-Jun-2001

C:Accession: G83735

R;Takami, H.; Nakasone, K.; Takai, Y.; Maeno, G.; Sasaki, R.; Masui, N.; Fuji, F.; Hira

Nucleic Acids Res. 28, 4317-4331, 2000

A:Title: Complete genome sequence of the alkaliphilic bacterium *Bacillus halodurans* and

A:Reference number: A83650; MID:20512582; PMID:11058132

A:Accession: G83735

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-458 <STO>

A:Cross-references: GB:AP001509; GB:BA000004; MID:g10173176; PIDN:BAB04406.1; GSPDB:GN00

A:Experimental source: strain C-125

C:Genetics:

A:Gene: BH0687

Query Match 3.5%; Score 8; DB 2; Length 458;  
Best Local Similarity 100.0%; Pred. No. 5.3;

Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 167 NPEOTKVL 174  
|||||

DB 291 NPEOTKVL 298

#### RESULT 13

RNA methyltransferase homolog yefA - *Bacillus subtilis*

C:Species: *Bacillus subtilis*

C>Date: 03-Dec-1997 #sequence\_revision 05-Dec-1997 #text\_change 20-Jun-2000

C:Accession: E69793

R;Kunst, F.; Ogasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Berier

C.; Bron, S.; Brouillet, S.; Hirsch, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.; Chd

A.; Ehlich, S.D.; Emerson, P.T.; Ertlan, K.D.; Errington, K.D.; Fabrev, C.; Ferrati, E.

Nature 390, 249-256, 1997

A:Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Galizzi, A.; Gall

lech, J.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Holsappel, S.; Hosono, S.; Hullo, M.F.

Koetter, P.; Koningsstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardinois,

A:Authors: Lauber, J.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, S.; Maueel

Y, M.; Ogawa, K.; Ogihara, A.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portetelle

Rieger, M.; Rivoita, C.; Rocha, E.; Roche, B.; Rose, M.; Sadate, Y.; Sato, T.; Scanlon,

A:Authors: Schleich, S.; Schreier, R.; Scoffone, F.; Sekiguchi, J.; Sekowska, A.; Serod

akuch, M.; Tanakoshi, A.; Tanaka, T.; Terpstra, P.; Tognoni, A.; Tosato, V.; Uchiyama,

T.; Winters, P.; Wipat, A.; Yamamoto, H.; Yamane, K.; Yasunori, K.; Yata, K.; Yoshida, K

A:Authors: Yoshikawa, H.F.; Zumbstein, E.; Yoshikawa, H.; Danchin, A.

A:Title: The complete genome sequence of the Gram-positive bacterium *Bacillus subtilis*.

A:Reference number: A69580; MID:98044033

A:Accession: E69793

A>Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-459 <KUN>

A:Cross-references: GB:299107; GB:AL009126; MID:g2632866; PIDN:CAB12493.1; PID:g26329

A:Experimental source: strain 168

C:Genetics:

A:Gene: yefA

C:Superfamily: hypothetical protein HI0333

Query Match 3.5%; Score 8; DB 2; Length 459;  
Best Local Similarity 100.0%; Pred. No. 5.3;

Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 167 NPEOTKVL 174  
|||||

DB 292 NPEOTKVL 299

#### RESULT 14

62K sucrose-binding protein homolog - garden pea

N:Alternate names: P54 protein

C:Species: *Pisum sativum* (garden pea)

C>Date: 23-Apr-1999 #sequence\_revision 23-Apr-1999 #text\_change 20-Jun-2000

C:Accession: T06459

R;Rodrigo II, M.I.

submitted to the EMBL Data Library, February 1997

A:Reference number: Z15693

A:Accession: T06459

A>Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-483 <ROD>

A:Cross-references: EMBL:Y11207; PIDN:CAA72090.1

A:Experimental source: var. Lincoln

C:Genetics:

A>Note: P54

C:Superfamily: glycine

Query Match 3.5%; Score 8; DB 2; Length 483;  
Best Local Similarity 100.0%; Pred. No. 5.6;

Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 151 AGTPMYLV 158  
|||||

DB 184 AGTPMYLV 191

#### RESULT 15

heat shock protein hscA [imported] - *Buchnera* sp. (strain APS)

C:Species: *Buchnera* sp.

C>Date: 02-Mar-2001 #sequence\_revision 02-Mar-2001 #text\_change 23-Mar-2001

C:Accession: A85000

R;Chienobu, S.; Matanabe, H.; Hattori, M.; Sakaki, Y.; Ishikawa, H.

Nature 407, 81-86, 2000

A:Title: Genome sequence of the endocellular bacterial symbiont of aphids *Buchnera* sp

A:Reference number: A84930; MID:20445173

A:Accession: A85000

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-611 <STO>

A:Cross-references: GB:AP000398; GSPDB:GN00144

A:Experimental source: strain APS

C:Genetics:

A:Gene: hscA; B0605

C:Superfamily: heat shock protein 70

Query Match 3.5%; Score 8; DB 2; Length 611;



Best Local Similarity 100.0%; Pred. No. 6.9;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 RYLLPSV 9  
|||||||  
Db 42 RYLLPSV 49

Search completed: August 28, 2002, 11:20:09  
Job time: 262 sec

**THIS PAGE BLANK (uspto)**

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 28, 2002, 11:18:47 ; Search time 21.78 Seconds

(without alignments)  
627.548 Million cell updates/sec

Title: US-09-853-526-4

Perfect score: 353  
Sequence: 1 MRVLLPSVVLGTAFTYVLA.....YVNTWVIGTLGLMTVRIKA 353

Scoring table:

OLIGO  
Gapop 60.0 , Gapext 60.0

Searched: 105224 seqs, 38719550 residues

Word size : 0

Total number of hits satisfying chosen parameters: 105224

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database : SwisProt\_40.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	353	100.0	353	1 PICE_HUMAN	Q9nq22 homo sapien
2	32	9.1	354	1 PICE_MOUSE	Q9dlie8 mus musculu
3	9	2.5	237	1 PLSC_HELPI	Q923n8 helicobacte
4	9	2.5	240	1 PLSC_HELPI	Q25903 helicobacte
5	8	2.3	83	1 V187_BPT3	P10302 bacterioph
6	8	2.3	83	1 V187_BPT3	P03788 bacterioph
7	8	2.3	396	1 V187_BPT3	P03788 bacterioph
8	8	2.3	611	1 HSCA_BUCAL	P57660 buchnera ap
9	7	2.0	57	1 YJDO_ECOLI	P58638 escherichia
10	7	2.0	194	1 PURI_LACCA	P35853 lactobacill
11	7	2.0	208	1 HBGF_CERAE	Q09118 cercopithe
12	7	2.0	208	1 HBGF_HUMAN	Q09075 homo sapien
13	7	2.0	208	1 HBGF_PIG	Q01580 sus scrofa
14	7	2.0	208	1 HBGF_PIG	Q06175 ratus norv
15	7	2.0	250	1 HBX4_MOUSE	P10284 mus musculu
16	7	2.0	250	1 PLSC_BORBU	Q09188 borrelia bu
17	7	2.0	250	1 PLSC_BORBU	Q03449 caenorhabd
18	7	2.0	251	1 HBX4_HUMAN	P17483 homo sapien
19	7	2.0	286	1 PROC_YEAST	P32265 saccharomy
20	7	2.0	308	1 PLSC_COCON	Q42670 cocos nucif
21	7	2.0	311	1 MPCR_YEAST	P23641 saccharomy
22	7	2.0	423	1 ENO_MERUA	Q60173 methanococ
23	7	2.0	445	1 XYL1_BACSU	P04748 bacillus su
24	7	2.0	516	1 GUX1_PCHAC	P13660 phanerocha
25	7	2.0	520	1 CMCH_NOCCLA	Q51080 nocardia la
26	7	2.0	532	1 GSI_NEUCRA	P38678 neurospora
27	7	2.0	677	1 Y593_TREPA	Q83632 treponema p
28	7	2.0	860	1 CH12_COCCIN	P54197 coccidioid
29	7	2.0	904	1 Y002_CAEEL	Q09228 caenorhabd
30	7	2.0	1018	1 ST31_MOUSE	Q09m1 mus musculu
31	7	2.0	1019	1 ST31_HUMAN	Q9bku1 homo sapien
32	7	2.0	1026	1 BGAL_STRTR	P23989 streptococ
33	7	2.0	1043	1 SYI_METHH	Q27428 methanobact

34	7	2.0	1044	1 SYI_METHH	P26499 methanobact
35	7	2.0	2485	1 PTND_HUMAN	Q12923 homo sapien
36	6	1.7	25	1 PCW1_PACGO	P82423 pachycondyl
37	6	1.7	25	1 PCW2_PACGO	P82424 pachycondyl
38	6	1.7	71	1 YAAA_BACSU	P05650 bacillus su
39	6	1.7	76	1 BB11_SCHCO	P78742 schizophyll
40	6	1.7	83	1 RL23_HALHA	Q06842 halobacteri
41	6	1.7	87	1 Y1WC_BPHH	P10435 bacterioph
42	6	1.7	88	1 PMRD_ECOLI	P37590 escherichia
43	6	1.7	88	1 Y4OL_RHISN	P55597 rhizobium s
44	6	1.7	99	1 RPOL_METVA	Q57832 methanococ
45	6	1.7	101	1 THST_THETS	Q9nbw4 thermozymon

#### ALIGNMENTS

RESULT	1	STANDARD	PRT	353 AA.
PLCE_HUMAN				
ID	PLCE_HUMAN			
AC	Q9N0G4			
DT	01-MAR-2002 (Rel. 41, Created)			
DT	01-MAR-2002 (Rel. 41, Last sequence update)			
DT	01-MAR-2002 (Rel. 41, Last annotation update)			
DE	1-acyl-sn-glycerol-3-phosphate acyltransferase epsilon (EC 2.3.1.51)			
DE	(1-AGPAT 5) (Lysophosphatidic acid			
DE	acyltransferase-epsilon) (LPAAT-epsilon) (1-acylglycerol-3-phosphate			
DE	O-acyltransferase 5).			
GN	AGPAT5.			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
OX	NCBI_TaxID=9606;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RA	Leung D.W.;			
RT	"Cloning and expression of LPAAT-epsilon.";			
RL	Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.			
RN	[2]			
RP	SEQUENCE FROM N.A.			
RA	Cohen D., Chumakov I., Blumenfeld M., Bougueleret L.;			
RL	Patent number WO9332644, 01-JUL-1999.			
RN	[3]			
RP	SEQUENCE FROM N.A.			
RA	TISSUE=amygdala;			
RX	MEDLINE=21154917; PubMed=11230166;			
RA	Wiemann S., Weill B., Wellenreuther R., Gassenhuber J., Glassl S.,			
RA	Ansorge W., Boecker M., Bloeker H., Bauersachs S., Blum H.,			
RA	Lauber J., Duesterhoeft A., Beyer A., Koehrer K., Strack N.,			
RA	Mewes H.-W., Othenwaelder B., Obermaier B., Tampe J., Heubner D.,			
RA	Wandutt R., Korn B., Klein M., Poustka A.;			
RT	"Towards a catalog of human genes and proteins: sequencing and			
RL	analysis of 500 novel complete protein coding human cDNAs.";			
RN	Genome Res. 11:422-435(2001).			
RP	[4]			
RP	SEQUENCE FROM N.A.			
RA	TISSUE=Placenta;			
RA	Isogai T., Ota T., Hayashi K., Sugiyama T., Otsuki T., Suzuki Y.,			
RA	Nishikawa T., Nagai K., Sugano S., Shiratori A., Sudo H.,			
RA	Wagatsashi M., Hosoi T., Kaku Y., Kodaira H., Kondo H., Sugawara M.,			
RA	Takahashi M., Chiba Y., Ishida S., Murakawa K., Ono Y., Takiguchi S.,			
RA	Watanabe S., Kimura K., Murakami K., Ishii S., Kawai Y., Saio K.,			
RA	Yamamoto J., Wakamatsu A., Nakamura Y., Nagahari K., Masuno Y.,			
RA	Niimiya K., Iwayanagi T.;			
RT	"NEO human cDNA sequencing project.";			
RL	Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.			
CC	- FUNCTION: CONVERTS LYSOPHOSPHATIDIC ACID (LPA) INTO PHOSPHATIDIC			
CC	ACID BY INCORPORATING ACYL MOIETY AT THE 2 POSITION (BY			
CC	SIMILARITY).			
CC	- CATALYTIC ACTIVITY: Acyl-CoA + 1-acyl-sn-glycerol 3-phosphate =			
CC	CoA + 1,2-diacyl-sn-glycerol 3-phosphate.			
CC	- PATHWAY: SECOND STEP IN DE NOVO PHOSPHOLIPID BIOSYNTHESIS.			
CC	- SUBCELLULAR LOCATION: Integral membrane protein (potential).			

CC -1- SIMILARITY: BELONGS TO THE 1-ACYL-SN-GLYCEROL-3-PHOSPHATE  
 CC ACYLTRANSFERASE FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL: AF375789; AAK54809.1; ALT\_INT.  
 DR EMBL: AL136587; CAB6522.1; ALT\_INIT.  
 DR EMBL: AK002072; BA92069.1; -  
 DR InterPro: IPR002123; Acyltransferase.  
 DR Pfam: PF01553; Acyltransferase; 1.  
 KM Phospholipid biosynthesis; Transferase; Acyltransferase;  
 KM Transmembrane.  
 FT TRANSMEM 7 29 POTENTIAL.  
 FT TRANSMEM 44 66 POTENTIAL.  
 FT TRANSMEM 334 351 POTENTIAL.  
 FT CONFLICT 145 145 L -> V (IN REF. 2).  
 SO SEQUENCE 353 AA; 40813 MW; A05B1FA246CE1B64 CRC64;

Query Match 100.0%; Score 353; DB 1; Length 353;  
 Best Local Similarity 100.0%; Pred. No. 0;  
 Matches 353; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MRYLPSVVLGTAPTYVLANGVRLSAPLPAFYQALDRLCYQVQSWLFFFEWVTG 60  
 DB 1 MRYLPSVVLGTAPTYVLANGVRLSAPLPAFYQALDRLCYQVQSWLFFFEWVTG 60  
 QY 61 VOILLYGDLPEKNENITYLANHSTVDIYADIAIRONALGHRYVLEKGLMPLPYGC 120  
 DB 61 VOILLYGDLPEKNENITYLANHSTVDIYADIAIRONALGHRYVLEKGLMPLPYGC 120  
 QY 121 YFAOHGIIYVRSKAKFENEKENRKLQSYVDAGTPMYLVEPEGRYNEQTKVLSASQAF 180  
 DB 121 YFAOHGIIYVRSKAKFENEKENRKLQSYVDAGTPMYLVEPEGRYNEQTKVLSASQAF 180  
 QY 181 AAOGLAVLVKHLTPRIKATPVAFDCKKNYLDAYDYTVYVEGDDGQRESPTMTFFL 240  
 DB 181 AAOGLAVLVKHLTPRIKATPVAFDCKKNYLDAYDYTVYVEGDDGQRESPTMTFFL 240  
 QY 241 CKEPKTIHIDRIKDVDEEOEHMRRLHERFEIKDKMLEFEYESPDPERRRRPFQGS 300  
 DB 241 CKEPKTIHIDRIKDVDEEOEHMRRLHERFEIKDKMLEFEYESPDPERRRRPFQGS 300  
 QY 301 VNSKLSIKTLPKMLISGLTASMLMTDAGRKLYVNTWYIGTLGCLMWTIKA 353  
 DB 301 VNSKLSIKTLPKMLISGLTASMLMTDAGRKLYVNTWYIGTLGCLMWTIKA 353

RESULT 2  
 PLSC\_MOUSE STANDARD; PRT; 354 AA.  
 AC 09D1E8;  
 DT 01-MAR-2002 (Rel. 41, Created)  
 DT 01-MAR-2002 (Rel. 41, Last sequence update)  
 DT 01-MAR-2002 (Rel. 41, Last annotation update)  
 DE 1-acyl-sn-glycerol-3-phosphate acyltransferase epsilon (EC 2.3.1.51)  
 DE (1-AGP acyltransferase 5) (1-AGPAT 5) (lysophosphatidic acid  
 DE acyltransferase-epsilon) (LPAT-epsilon) (1-acylglycerol-3-phosphate  
 DE O-acyltransferase 5).  
 OS AGPAT5.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=C57BL/6J; TISSUE=Embryo;

RX MEDLINE=21085660; PubMed=11217851;  
 RA Kawai J., Shingawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,  
 RA Arakawa T., Hara A., Fukunishi Y., Kono H., Adachi J., Fukuda S.,  
 RA Alzawa K., Iwama M., Nishi K., Kiyosawa H., Kondo S., Yamana K. I.,  
 RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,  
 RA Kadota K., Matsuda H. A., Ashburner M., Batalov S., Casavant T.,  
 RA Fleischmann W., Gaasterland T., Glasi C., King B., Kochiwa H.,  
 RA Kiehl P., Lewis S., Matsuo Y., Nikaide I., Pesole G., Quackenbush J.,  
 RA Schirral L. M., Staubli F., Suzuki R., Tomita M., Wagner L., Mashio T.,  
 RA Sakai K., Okido T., Furuno M., Aono H., Balderelli R., Barsh G.,  
 RA Blake J., Boffelli D., Bojunga N., Carrincci P., de Bonaldo M. F.,  
 RA Brownstein M. J., Bult C., Fletcher C., Fujita M., Gariboldi M.,  
 RA Gustinich S., Hill D., Hofmann M., Hume D. A., Kamiya M., Lee N. H.,  
 RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,  
 RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,  
 RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,  
 RA Suzuki H., Toyooka K., Wang K. H., Weltz C., Whitaker C., Wilmink L.,  
 RA Wyszaw-Boris A., Yoshida K., Hasegawa Y., Kawai H., Kohsaki S.,  
 RA Hayashizaki Y.,  
 RA "Functional annotation of a full-length mouse cDNA collection."  
 RL Nature 409:685-690(2001).  
 CC -1- FUNCTION: CONVERTS LYSOPHOSPHATIDIC ACID (LPA) INTO PHOSPHATIDIC  
 CC ACID BY INCORPORATING ACYL MOIETY AT THE 2 POSITION (BY  
 CC SIMILARITY).  
 CC -1- CATALYTIC ACTIVITY: Acyl-CoA + 1-acyl-sn-glycerol 3-phosphate =  
 CC CoA + 1,2-diacyl-sn-glycerol 3-phosphate.  
 CC -1- PATHWAY: SECOND STEP IN DE NOVO PHOSPHOLIPID BIOSYNTHESIS.  
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein (Potential).  
 CC -1- SIMILARITY: BELONGS TO THE 1-ACYL-SN-GLYCEROL-3-PHOSPHATE  
 CC ACYLTRANSFERASE FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL: AK003649; BAB22915.1; -  
 DR MGD: MGI:1915880; Agpat5.  
 DR InterPro: IPR002123; Acyltransferase.  
 DR Pfam: PF01553; Acyltransferase; 1.  
 KM Phospholipid biosynthesis; Transferase; Acyltransferase;  
 KM Transmembrane.  
 FT TRANSMEM 7 29 POTENTIAL.  
 FT TRANSMEM 44 66 POTENTIAL.  
 FT TRANSMEM 335 352 POTENTIAL.  
 SO SEQUENCE 354 AA; 40943 MW; 447BEF924B91E800 CRC64;

Query Match 9.1%; Score 32; DB 1; Length 354;  
 Best Local Similarity 100.0%; Pred. No. 1.3e-24;  
 Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 174 LSASQAFAPQAGLAVLKHLTPRIKATPVAFD 205  
 DB 174 LSASQAFAPQAGLAVLKHLTPRIKATPVAFD 205  
 RESULT 3  
 PLSC\_HELPJ STANDARD; PRT; 237 AA.  
 AC 09ZJN8;  
 DT 16-OCT-2001 (Rel. 40, Created)  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE 1-acyl-sn-glycerol-3-phosphate acyltransferase (EC 2.3.1.51) (1-AGP  
 DE acyltransferase) (1-AGPAT) (lysophosphatidic acid acyltransferase)  
 DE (LPATV).  
 OS PLSC OR JHP1267.  
 OS Helicobacter pylori j99 (Campylobacter pylori j99).  
 OC Bacteria; Proteobacteria; epsilon subdivision; Helicobacter group;

```

OC Helicobacter.
OX NCBI_TaxID=85963;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=99120557; PubMed=9923682;
RA Alm R.A., Ling L.-S.L., Moir D.T., King B.L., Brown E.D., Doig P.C.,
RA Smith D.R., Noonan B., Guild B.C., deDonge B.L., Carmel G.,
RA Tummino P.J., Caruso A., Uria-Nickelsen M., Mills D.M., Ives C.,
RA Gibson R., Merberg D., Mills S.D., Jiang Q., Taylor D.E., Vovis G.F.,
RA Trust T.J.;
RT "Genomic sequence comparison of two unrelated isolates of the human
RT gastric pathogen Helicobacter pylori."
RL Nature 397:176-180(1999).
CC -1- FUNCTION: CONVERTS LYSOPHOSPHATIDIC ACID (LPA) INTO PHOSPHATIDIC
CC ACID BY INCORPORATING ACYL MOIETY AT THE 2 POSITION.
CC -1- CATALYTIC ACTIVITY: Acyl-CoA + 1-acyl-sn-glycerol 3-phosphate =
CC COA + 1,2-diacyl-sn-glycerol 3-phosphate.
CC -1- PATHWAY: SECOND STEP IN DE NOVO PHOSPHOLIPID BIOSYNTHESIS.
CC -1- SUBCELLULAR LOCATION: Inner membrane-associated (Potential).
CC -1- SIMILARITY: BELONGS TO THE 1-ACYL-SN-GLYCEROL-3-PHOSPHATE
CC ACYLTRANSFERASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AE001550; AAD06852.1; -
DR InterPro: IPR002123; Acyltransferase.
DR Pfam: PF01553; Acyltransferase; 1.
KW Phospholipid biosynthesis; Transferase; Acyltransferase;
KW Inner membrane; Complete proteome.
SQ SEQUENCE 237 AA; 27188 MW; E10F517D42A1731F CRC64;

Query Match 2.5%; Score 9; DB 1; Length 237;
Best Local Similarity 100.0%; Pred. No. 0.2;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 157 LVTFPEGTR 165
DB 142 LVTFPEGTR 150

RESULT 4
PLSC_HELPY STANDARD; PRT; 240 AA.
AC 025903;
DT 15-DEC-1998 (rel. 37, Created)
DT 15-DEC-1998 (rel. 37, Last sequence update)
DT 16-OCT-2001 (rel. 40, Last annotation update)
DE 1-acyl-sn-glycerol-3-phosphate acyltransferase (EC 2.3.1.51) (1-AGP
DE acyltransferase) (1-AGPAT) (lysophosphatidic acid acyltransferase)
DE (LPAAT).
DE PLSC OR HP1348.
OS Helicobacter pylori (Campylobacter pylori).
OC Bacteria; Proteobacteria; epsilon subdivision; Helicobacter group;
OC Helicobacter.
OX NCBI_TaxID=210;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=26695 / ATCC 700392;
RX MEDLINE=97394467; PubMed=9252185;
RA Tomb J.-F., White O., Kierlavage A.R., Clayton R.A., Sutton G.G.,
RA Fleischmann R.D., Ketchum K.A., Klenk H.-P., Gill S., Dougherty B.A.,
RA Nelson K., Quackenbush J., Zhou L., Kirkness E.F., Peterson S.,
RA Loftus B., Richardson D., Dodson R., Khalak H.G., Glodek A.,
RA McInerney K., Fitzgerald L.M., Lee N., Adams M.D., Hickey E.K.,
RA Berg D.E., Gocayne J.D., Utterback T.R., Peterson J.D., Kelley J.M.,
RA Cotton M.D., Meldrum J.M., Fujii C., Bowman C., Wathey L., Wallin E.,

```

```

RA Hayes W.S., Borodovsky M., Karp P.D., Smith H.O., Fraser C.M.,
RA Venter J.C.;
RT "The complete genome sequence of the gastric pathogen Helicobacter
RT pylori."
RL Nature 388:539-547(1997).
CC -1- FUNCTION: CONVERTS LYSOPHOSPHATIDIC ACID (LPA) INTO PHOSPHATIDIC
CC ACID BY INCORPORATING ACYL MOIETY AT THE 2 POSITION.
CC -1- CATALYTIC ACTIVITY: Acyl-CoA + 1-acyl-sn-glycerol 3-phosphate =
CC COA + 1,2-diacyl-sn-glycerol 3-phosphate.
CC -1- PATHWAY: SECOND STEP IN DE NOVO PHOSPHOLIPID BIOSYNTHESIS.
CC -1- SUBCELLULAR LOCATION: Inner membrane-associated (Potential).
CC -1- SIMILARITY: BELONGS TO THE 1-ACYL-SN-GLYCEROL-3-PHOSPHATE
CC ACYLTRANSFERASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AE000636; AAD08393.1; -
DR TIGR: HP1348; -
DR InterPro: IPR002123; Acyltransferase.
DR Pfam: PF01553; Acyltransferase; 1.
KW Phospholipid biosynthesis; Transferase; Acyltransferase;
KW Inner membrane; Complete proteome.
SQ SEQUENCE 240 AA; 27745 MW; 22BD5D0EB190BBD CRC64;

Query Match 2.5%; Score 9; DB 1; Length 240;
Best Local Similarity 100.0%; Pred. No. 0.2;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 157 LVTFPEGTR 165
DB 142 LVTFPEGTR 150

RESULT 5
V187_BP73 STANDARD; PRT; 83 AA.
AC P10302;
DT 01-MAR-1989 (rel. 10, Created)
DT 01-MAR-1989 (rel. 10, Last sequence update)
DT 01-MAR-1989 (rel. 10, Last annotation update)
DE Gene 18.7 protein.
DE 18.7.
OS Bacteriophage T3.
OC Viruses; dsDNA viruses, no RNA stage; Caudovirales; Podoviridae;
OC T7-like phages.
OX NCBI_TaxID=10759;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=86209997; PubMed=3010556;
RA Yamada M., Fujisawa H., Kato H., Hamada K., Minagawa T.;
RT "Cloning and sequencing of the genetic right end of bacteriophage T3
RT DNA."
RL Virology 151:350-361(1986).
RN [2]
RP ERRATUM.
RA Yamada M., Fujisawa H., Kato H., Hamada K., Minagawa T.;
RL Virology 154:246-246(1986).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----

```

DR EMBL: M14784; AAA92527.1; -.  
 DR PIR: E23476; W8BPT3.  
 SQ SEQUENCE 83 AA; 9393 MW; 63128984CBAF0531 CRC64;

Query Match 2.3%; Score 8; DB 1; Length 83;  
 Best Local Similarity 100.0%; Pred. No. 0.86;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 305 LSIRKTLTP 312  
 |||||||  
 Db 22 LSIRKTLTP 29

RESULT 6  
 V187\_BPT7 STANDARD; PRT; 83 AA.  
 AC P03788;  
 DT 21-JUL-1986 (Rel. 01, Last sequence update)  
 DT 21-JUL-1986 (Rel. 01, Last sequence update)  
 DT 01-MAR-1989 (Rel. 10, Last annotation update)  
 DE Gene 18.7 protein.  
 GN 18.7.  
 OS Bacteriophage T7.  
 OC Viruses; dsDNA viruses, no RNA stage; Caudovirales; Podoviridae;  
 CC T7-like phages.  
 OX NCBI\_TaxID=10760;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=83241725; PubMed=6864790;  
 RA Dunn J.J., Studier F.W.;  
 RT "Complete nucleotide sequence of bacteriophage T7 DNA and the  
 RT locations of T7 genetic elements."  
 RT J. Mol. Biol. 166:477-535(1983).  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL: V01146; CAA24439.1; -.  
 DR PIR: A04413; W8BPG7.  
 DR PIR: S42337; S42337.  
 SQ SEQUENCE 83 AA; 9326 MW; B3DB055DE79A53C5 CRC64;

Query Match 2.3%; Score 8; DB 1; Length 83;  
 Best Local Similarity 100.0%; Pred. No. 0.86;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 305 LSIRKTLTP 312  
 |||||||  
 Db 22 LSIRKTLTP 29

RESULT 7  
 YD18\_YEAST STANDARD; PRT; 396 AA.  
 ID YD18\_YEAST  
 AC Q12185;  
 DT 15-JUL-1999 (Rel. 38, Created)  
 DT 15-JUL-1999 (Rel. 38, Last sequence update)  
 DT 15-JUL-1999 (Rel. 38, Last annotation update)  
 DE Hypothetical 45.9 kDa protein in KCS1-GCV1 intergenic region.  
 GN YD018C OR Y09335.04C OR PZF396.  
 OS Saccharomyces cerevisiae (Baker's yeast).  
 CC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;  
 CC Saccharomycetales; Saccharomycetaceae; Saccharomyces.  
 OX NCBI\_TaxID=4932;  
 CC [1]  
 RN RP  
 SEQUENCE FROM N.A.

RC STRAIN-S288C / AB972.  
 RA Dedman K., Brown D., Hamlyn N., Bowman S., Barrell B.G.,  
 RA Rajandream M.A.;  
 RL Submitted (MAY-1995) to the EMBL/GenBank/DBJ databases.

RN [2]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=97051598; PubMed=8896275;  
 RA Elde L.G., Sander C., Prydz H.;  
 RT "Sequencing and analysis of a 35.4 kb region on the right arm of  
 RT chromosome IV from Saccharomyces cerevisiae reveal 23 open reading  
 RT frames."  
 RL Yeast 12:1085-1090(1996).  
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein (potential).  
 CC -1- SIMILARITY: BELONGS TO THE 1-ACYL-SN-GLYCEROL-3-PHOSPHATE  
 CC ACYLTRANSFERASE FAMILY.  
 CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----

CC EMBL: Z74314; CAA98838.1; -.  
 DR EMBL: X95966; CAA65210.1; -.  
 DR EMBL: Z49770; CAA89843.1; -.  
 DR SGD: S0002425; YDR018C.  
 DR InterPro: IPR002123; Acyltransferase.  
 DR Pfam: PF01553; Acyltransferase; 1.  
 KW Hypothetical protein; phospholipid biosynthesis; Transferase;  
 KW Acyltransferase; Transmembrane.  
 FT TRANSMEM 27 47  
 FT TRANSMEM 69 89 POTENTIAL.  
 FT TRANSMEM 123 143 POTENTIAL.  
 FT TRANSMEM 372 392 POTENTIAL.  
 SQ SEQUENCE 396 AA; 45938 MW; 9F55AB56C8BD44DD CRC64;

Query Match 2.3%; Score 8; DB 1; Length 396;  
 Best Local Similarity 100.0%; Pred. No. 3.2;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 211 LDATYDVT 218  
 |||||||  
 Db 265 LDATYDVT 272

RESULT 8  
 HSCA\_BUCAI STANDARD; PRT; 611 AA.  
 ID HSCA\_BUCAI  
 AC P57660;  
 DT 16-OCT-2001 (Rel. 40, Created)  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Chapterone protein hscA homolog.  
 GN HSCA OR BU605.  
 OS Buchnera aphidicola (subsp. Acyrthosiphon pisum) (Acyrthosiphon pisum  
 OS symbiotic bacterium).  
 CC Bacteria; Proteobacteria; gamma subdivision; Buchnera.  
 OX NCBI\_TaxID=118099;  
 CC [1]  
 RN RP  
 SEQUENCE FROM N.A.  
 RC STRAIN-TOKYO 1998;  
 RX MEDLINE=20445173; PubMed=10993077;  
 RA Shigenobu S., Watanabe H., Hattori M., Sakaki Y., Ishikawa H.;  
 RT "Genome sequence of the endocellular bacterial symbiont of aphids  
 RT Buchnera sp. Aps."  
 RL Nature 407:81-86(2000).  
 CC -1- FUNCTION: PROBABLE CHAPERONE. HAS A LOW INTRINSIC ATPASE ACTIVITY  
 CC WHICH IS MARKEDLY STIMULATED BY HSCB (BY SIMILARITY).  
 CC -1- SIMILARITY: BELONGS TO THE HEAT SHOCK PROTEIN 70 FAMILY.  
 CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; AP001119; BAB13289.1; -  
 DR Interpro: IPR001023; HSP70.  
 DR Pfam: PF00012; HSP70.1.  
 DR PRINTS: PR00301; HEATSHOCK70.  
 DR PROSITE: PS00297; HSP70.1; 1.  
 DR PROSITE: PS00329; HSP70.2; 1.  
 DR PROSITE: PS01036; HSP70.3; 1.  
 KW Chaperone: ATP-binding: Complete proteome.  
 SQ SEQUENCE 611 AA: 69837 MW: 32567C53073082A CRC64;

Query Match 2.3%; Score 8; DB 1; Length 611;  
 Best Local Similarity 100.0%; Pred. No. 4.6;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 RYLPSV 9  
 Db 42 RYLPSV 49

RESULT 9  
 YJDO\_ECOLI STANDARD: PRT; 57 AA:  
 AC P58038;  
 DT 16-OCT-2001 (Rel. 40, Created)  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DT 01-MAR-2002 (Rel. 41, Last annotation update)  
 DE Hypothetical protein yjdo.  
 GN YJDO OR B4128.1 OR Z5731 OR EC55110.1.  
 OS Escherichia coli, and  
 OS Escherichia coli O157:H7.  
 OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;  
 OC Escherichia.  
 ON NCBI\_Taxid=562, 83334;  
 RX SEQUENCE FROM N.A.  
 RC STRAIN-K12 / MG1655;  
 RA Blatner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,  
 RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,  
 RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,  
 RA Mau B., Shao Y.;  
 RT "The complete genome sequence of Escherichia coli K-12."  
 RL Science 277:1453-1474(1997).  
 RN [12]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-O157:H7 / EDL933 / ATCC 700927;  
 RX MEDLINE-21074935; PubMed-11206551;  
 RA Perna N.T., Plunkett G. III, Burland V., Mau B., Glasner J.D.,  
 RA Rose D.J., Mayhew G.F., Evans P.S., Gregor J., Kirkpatrick H.A.,  
 RA Postel G., Hackett J., Link S., Boutin A., Shao Y., Miller L.,  
 RA Grobeck E.J., Davis N.W., Lim A., Dimalanta E.T., Potamousis K.,  
 RA Apodaca J., Manthariman T.S., Lin J., Yen G., Schwartz D.C.,  
 RA Welch R.A., Blatner F.R.;  
 RT "Genome sequence of enterohemorrhagic Escherichia coli O157:H7."  
 RL Nature 409:529-533(2001).  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-O157:H7 / RMD 0509952;  
 RX MEDLINE-21156231; PubMed-11258796;  
 RA Hayashi T., Makino K., Ohnishi M., Kurokawa K., Ishii K., Yokoyama K.,  
 RA Han C.-G., Ohtsubo E., Nakayama K., Murata T., Tanaka M., Tobe T.,  
 RA Iida T., Takami H., Honda T., Sasaki K., Ogasawara N., Yasunaga T.,  
 RA Kihara S., Shiba T., Hattori M., Shinagawa H.;  
 RT "Complete genome sequence of enterohemorrhagic Escherichia coli

RT O157:H7 and genomic comparison with a laboratory strain K-12."  
 RL DNA Res. 8:11-22(2001).  
 CC -1 SUBCELLULAR LOCATION: Integral membrane protein (Potential).  
 CC -1 SIMILARITY: STRONG. TO E.COLI YDCX.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; AE000485; -; NOT\_ANNOTATED\_CDS.  
 DR EMBL; AE003646; AAG59328.1; -;  
 DR EMBL; AP002568; -; NOT\_ANNOTATED\_CDS.  
 DR EcoGene: EG14342; yjdo.  
 KW Hypothetical protein; Transmembrane; Complete proteome.  
 FT TRANSMEM 7 57  
 FT TRANSMEM 37 57 POTENTIAL.  
 SQ SEQUENCE 57 AA: 6555 MW: A4670A19500F75D6 CRC64;

Query Match 2.0%; Score 7; DB 1; Length 57;  
 Best Local Similarity 100.0%; Pred. No. 6.5;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 25 RLISAF 31  
 Db 33 RLISAF 39

RESULT 10  
 PURI\_LACCA STANDARD: PRT; 194 AA:  
 AC P35853;  
 DT 01-JUN-1994 (Rel. 29, Created)  
 DT 01-FEB-1996 (Rel. 33, Last sequence update)  
 DT 30-MAY-2000 (Rel. 39, Last annotation update)  
 DE Amidophosphoribosyltransferase precursor (EC 2.4.2.14) (Glutamine  
 DE phosphoribosylpyrophosphate amidotransferase) (ATASE) (GPATASE)  
 DE (Fragment).  
 DE PURF.  
 GN Lactobacillus casei.  
 OS Bacteria; Firmicutes; Bacillus/Clostridium group; Lactobacillaceae;  
 OC Lactobacillus.  
 ON NCBI\_Taxid=1562;  
 RX SEQUENCE FROM N.A.  
 RA MEDLINE-93012962; PubMed-1398079;  
 RA Gu Z.-M., Martindale D.W., Lee B.H.;  
 RT "Isolation and complete sequence of the purL gene encoding PGAM  
 RT synthase II in Lactobacillus casei."  
 RL Gene 119:123-126(1992).  
 RN [12]  
 RP ERRATUM.  
 RX MEDLINE-94040790; PubMed-8224889;  
 RA Gu Z.-M., Martindale D.W., Lee B.H.;  
 RL Gene 133:147-147(1993).  
 CC -1 CATALYTIC ACTIVITY: 5-phospho-beta-D-ribose 1-diphosphate +  
 CC L-glutamate = L-glutamine + 5-phospho-alpha-D-ribose 1-diphosphate  
 CC + H(2)O.  
 CC -1 PATHWAY: FIRST STEP IN DE NOVO PURINE BIOSYNTHESIS.  
 CC -1 SIMILARITY: THE GATASE DOMAIN BELONGS TO TYPE-2 GLUTAMINE  
 CC AMIDOTRANSFERASES.  
 CC -1 PURINE/PYRIMIDINE PHOSPHORIBOSYLTRANSFERASE FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC -----

DR EMBL: M65265; AAC36948.1; -.

DR PIR: PC1136; PC1136.

DR HSSP: P00497; 1A00.

DR MEROPS: C44.001; -.

DR InterPro: IPP000583; GATase\_2.

DR InterPro: IP020375; Pur\_pyr\_pr\_transf.

DR Pfam: PF00310; GATase\_2; 1.

DR PROSITE: PS00103; PUR\_PYR\_PR\_TRANSFER; PARTIAL.

DR PROSITE: PS00443; GATASE\_TYPE\_II; 1.

KW Purine biosynthesis; Transferase; Glycosyltransferase.

FT PROPEP 1 11

FT CHAIN 12 >194

FT ACT\_SITE 12 12

FT NON\_TER 194 194

FT SEQUENCE 194 AA; 21144 MW; 4A788BC5365D5EC CRC64;

Query Match 2.08; Score 7; DB 1; Length 194; Best Local Similarity 100.0%; Pred. No. 18; Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 100 ALGHVRY 106

DB 79 ALGHVRY 85

RESULT 11

HBGF\_CERAE STANDARD; PRT; 208 AA.

AC 009118:

DT 01-FEB-1996 (Rel. 33, Created)

DT 01-FEB-1996 (Rel. 33, Last sequence update)

DT 16-OCT-2001 (Rel. 40, Last annotation update)

DE Heparin-binding EGF-like growth factor precursor (HB-EGF) (HBEGF)

DE (Diphtheria toxin receptor) (DT-R).

GN DTR OR HEGFL.

OS Cercopithecus aethiops (Green monkey) (Grivet).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea; Cercopithecidae; Cercopithecus.

NCBI\_Taxid:9534;

[1]

RP SEQUENCE FROM N.A.

RA MEDLINE-92298386; PubMed-1606612;

RA Naglich J.G.; Metherall J.E.; Russel D.W.; Eldels L.;

RT "Expression cloning of a diphtheria toxin receptor: identity with a heparin-binding EGF-like growth factor precursor.";

RL Cell 69:1051-1061(1992).

[2]

RN TOXIN-BINDING DOMAIN.

RX MEDLINE-95126975; PubMed-7826391;

RA Hooper K.P.; Eldels L.;

RT "Localization of a critical diphtheria toxin-binding domain to the C-terminus of the mature heparin-binding EGF-like growth factor region of the diphtheria toxin receptor.";

RL Biochem. Biophys. Res. Commun. 206:710-717(1995).

CC -1- FUNCTION: MAY BE INVOLVED IN MACROPHAGE-MEDIATED CELLULAR PROLIFERATION. IT IS MITOGENIC FOR FIBROBLASTS AND SMOOTH MUSCLE BUT NOT ENDOTHELIAL CELLS. IT IS ABLE TO BIND EGF RECEPTORS WITH HIGHER AFFINITY THAN EGF ITSELF AND IS A FAR MORE POTENT MITOGEN FOR SMOOTH MUSCLE CELLS THAN EGF (BY SIMILARITY).

CC -1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN. MATURE HB-EGF IS RELEASED INTO THE EXTRACELLULAR SPACE AND PROBABLY BINDS TO A RECEPTOR (BY SIMILARITY).

CC -1- PTM: O-GLYCOSYLATED (BY SIMILARITY).

CC -1- SIMILARITY: CONTAINS 1 EGF-LIKE DOMAIN.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC -----

DR EMBL: M93012; -; NOT\_ANNOTATED\_CDS.

DR HSSP: Q99075; 1XDT.

DR InterPro: IPR000561; EGF-like.

DR Pfam: PF00008; EGF; 1.

DR SMART: SM00181; EGF; 1.

DR PROSITE: PS00022; EGF\_1; 1.

DR PROSITE: PS0186; EGF\_2; 1.

KW Growth factor; Heparin-binding; Signal; EGF-like domain; Transmembrane; Glycoprotein; Receptor.

FT SIGNAL 1 19

FT PROPEP 20 62

FT CHAIN 63 148

FT PROPEP 149 208

FT DOMAIN 20 160

FT TRANSMEM 161 184

FT DOMAIN 185 208

FT DOMAIN 104 144

FT CARBOHYD 75 75

FT CARBOHYD 85 85

FT DISULFID 108 121

FT DISULFID 116 132

FT DISULFID 134 143

FT SEQUENCE 208 AA; 22985 MW; 8D108289A0485AE9 CRC64;

Query Match 2.08; Score 7; DB 1; Length 208; Best Local Similarity 100.0%; Pred. No. 19; Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 4 LIPSVYL 10

DB 3 LIPSVYL 9

RESULT 12

HBGF\_HUMAN STANDARD; PRT; 208 AA.

AC 099075:

DT 01-JUN-1994 (Rel. 29, Created)

DT 01-JUN-1994 (Rel. 29, Last sequence update)

DT 16-OCT-2001 (Rel. 40, Last annotation update)

DE Heparin-binding EGF-like growth factor precursor (HB-EGF) (HBEGF)

DE (Diphtheria toxin receptor) (DT-R).

GN DTR OR HEGFL.

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiindae; Homo.

NCBI\_Taxid:9606;

[1]

RP SEQUENCE FROM N.A. AND SEQUENCE OF 73-93.

RX TISSUE-Macrophage;

RA MEDLINE-91157008; PubMed-1840698;

RA Higashiyama S.; Abraham J.A.; Miller J.; Fiddes J.C.; Klagsbrun M.;

RT "A heparin-binding growth factor secreted by macrophage-like cells that is related to EGF.";

RL Science 251:936-939(1991).

[2]

RN SEQUENCE FROM N.A.

RA Klimmerly W.; Bondoc M.; Cheng J.; Connolly K.S.; Gunning K.M.; Kadner K.; Miguel T.; Miller C.; Ptluck S.; Pollard M.; Rojeski H.; Subramanian S.; Martin C.H.;

RL Submitted (Apr-1998) to the EMBL/GenBank/DBJ databases.

[3]

RP SEQUENCE OF 63-141 AND 143-148, AND CARBOHYDRATE-LINKAGE SITES.

RX TISSUE-Histiocytic lymphoma;

RA MEDLINE-92210596; PubMed-1556128;

RA Higashiyama S.; Lau K.; Besner G.E.; Abraham J.A.; Klagsbrun M.;



RT "Structure of heparin-binding EGF-like growth factor. Multiple forms,  
RT primary structure, and glycosylation of the mature protein.";  
RL J. Biol. Chem. 267:6205-6212(1992).  
RN [4]  
RX TOXIN-BINDING DOMAIN.  
RX MEDLINE-95138082; PubMed-7836353;  
RA Mitamura T., Higashiyama S., Taniguchi N., Klagsbrun M., Mekada E.;  
RT "Diphtheria toxin binds to the epidermal growth factor (EGF)-like  
RT domain of human heparin-binding EGF-like growth factor/diphtheria  
RT toxin receptor and inhibits specifically its mitogenic activity.";  
RL J. Biol. Chem. 270:1015-1019(1995).  
RN [5]  
RP X-RAY CRYSTALLOGRAPHY (2.65 ANGSTROMS) OF 73-147 IN COMPLEX WITH TOX.  
RX MEDLINE-98324089; PubMed-9659904;  
RA Louis G.V., Yang W., Bowman M.E., Choe S.;  
RT "Crystal structure of the complex of diphtheria toxin with an  
RT extracellular fragment of its receptor.";  
RL Mol. Cell 1:67-78(1997).  
CC -1- FUNCTION: MAY BE INVOLVED IN MACROPHAGE-MEDIATED CELLULAR  
CC PROLIFERATION. IT IS MITOGENIC FOR FIBROBLASTS AND SMOOTH MUSCLE  
CC BUT NOT ENDOTHELIAL CELLS. IT IS ABLE TO BIND EGF RECEPTORS WITH  
CC HIGHER AFFINITY THAN EGF ITSELF AND IS A FAR MORE POTENT MITOGEN  
CC FOR SMOOTH MUSCLE CELLS THAN EGF.  
CC -1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN. MATURE HB-EGF IS  
CC RELEASED INTO THE EXTRACELLULAR SPACE AND PROBABLY BINDS TO A  
CC RECEPTOR.  
CC -1- PTM: SEVERAL N-TERMINI HAVE BEEN IDENTIFIED BY DIRECT SEQUENCING.  
CC THE FORMS WITH N-TERMINI 63, 73 AND 74 HAVE BEEN TESTED AND FOUND  
CC TO BE BIOLOGICALLY ACTIVE.  
CC -1- PTM: O-GLYCAN ATTACHMENT SITES WERE DETERMINED BY EDMAN  
CC DEGRADATION. O-GLYCANASE DIGEST SUGGESTS MUCIN-TYPE GLYCOSYLATION  
CC (DONE IN HB-EGF PURIFIED FROM HISTIOCYTIC LYMPHOMA CELL LINE  
CC U-937).  
CC -1- SIMILARITY: CONTAINS 1 EGF-LIKE DOMAIN.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sdb.ch/announce/>  
CC or send an email to [license@sdb-sib.ch](mailto:license@sdb-sib.ch)).  
CC -----  
CC EMBL: M60278; AAA35956.1; -  
CC EMBL: AC004634; AAC15470.1; -  
CC PIR: A38432; A38432.  
CC PDB: 1XDT; 25-FEB-98.  
CC MIM: 126150; -  
CC InterPro: IPR000561; EGF-like.  
CC DR Pfam: PF00008; EGF\_1.  
CC DR SMART: SM00181; EGF\_1.  
CC DR PROSITE: PS00022; EGF\_1; 1.  
CC DR PROSITE: PS01186; EGF\_2; 1.  
CC KW Growth factor; Heparin-binding; Signal; EGF-like domain;  
CC Transmembrane; Glycoprotein; Receptor; 3D-structure.  
CC FT SIGNAL 1 19  
CC FT PROPEP 20 62  
CC FT CHAIN 63 148  
CC FT PROPEP 149 208  
CC FT DOMAIN 20 160  
CC FT TRANSMEM 161 184  
CC FT TRANSMEM 185 208  
CC FT DOMAIN 104 144  
CC FT DOMAIN 108 121  
CC FT DISULFID 116 132  
CC FT DISULFID 134 143  
CC FT DISULFID 116 132  
CC FT CARBOHYD 85 75  
CC FT CARBOHYD 85 75  
CC SEQUENCE 208 AA; 23067 MW; 2CA3C9D1B8291B51 CRG64;  
Query Match 2.0%; Score 7; DB 1; Length 208;  
Best Local Similarity 100.0%; Pred. No. 19;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 4 LIPSVYL 10  
DB 3 LIPSVYL 9  
RESULT 13  
ID HBGF\_PIG STANDARD; PRT; 208 AA.  
AC 001580;  
DT 01-FEB-1996 (Rel. 33, Created)  
DT 30-MAY-2000 (Rel. 39, Last sequence update)  
DE 16-OCT-2001 (Rel. 40, Last annotation update)  
DE Heparin-binding EGF-like growth factor precursor (HB-EGF) (HBEGF).  
GN DTR OR HBGL.  
OS Sus scrofa (Pig).  
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.  
OX NCBI\_TaxID=9823;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Pascall J.C.;  
RL Submitted (NOV-1997) to the EMBL/Genbank/DBJ databases.  
RN [2]  
RP SEQUENCE OF 99-182 FROM N.A.  
RC STRAIN=LARGE WHITE; TISSUE=Heart;  
RX MEDLINE-93075016; PubMed-1445231;  
RA Vaughan T.J., Pascall J.C., Brown K.D.;  
RT "Tissue distribution of mRNA for heparin-binding epidermal growth  
RT factor";  
RL Biochem. J. 287:681-684(1992).  
CC -1- FUNCTION: MAY BE INVOLVED IN MACROPHAGE-MEDIATED CELLULAR  
CC PROLIFERATION. IT IS MITOGENIC FOR FIBROBLASTS AND SMOOTH MUSCLE  
CC BUT NOT ENDOTHELIAL CELLS. IT IS ABLE TO BIND EGF RECEPTORS WITH  
CC HIGHER AFFINITY THAN EGF ITSELF AND IS A FAR MORE POTENT MITOGEN  
CC FOR SMOOTH MUSCLE CELLS THAN EGF (BY SIMILARITY).  
CC -1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN. MATURE HB-EGF IS  
CC RELEASED INTO THE EXTRACELLULAR SPACE AND PROBABLY BINDS TO A  
CC RECEPTOR (BY SIMILARITY).  
CC -1- TISSUE SPECIFICITY: MACROPHAGE, MIDBRAIN, CEREBELLUM,  
CC HYPOTHALAMUS, CEREBRAL CORTEX, BUTYBORETHRAL GLAND, LUNG, HEART  
CC VENTRICLE, KIDNEY, SKIN, PROSTATE, SEMINAL VESICLE, TESTIS; AT LOW  
CC LEVELS IN LYMPH NODE, THYMUS, SPLEEN, NOT DETECTED IN PITUITARY,  
CC OLFACTORY BULB, THYROID, DUODENUM, PANCREAS, LIVER, SUBMAXILLARY  
CC GLAND.  
CC -1- PTM: O-GLYCOSYLATED (PROBABLE).  
CC -1- SIMILARITY: CONTAINS 1 EGF-LIKE DOMAIN.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sdb.ch/announce/>  
CC or send an email to [license@sdb-sib.ch](mailto:license@sdb-sib.ch)).  
CC -----  
CC EMBL: Y15731; CAA75740.1; -  
CC EMBL: X67295; CAA47709.1; -  
CC HSSP: Q99075; 1XDT.  
CC InterPro: IPR000561; EGF-like.  
CC DR InterPro: IPR001336; EGF\_1.  
CC DR Pfam: PF00008; EGF\_1.  
CC DR PRINTS: PR00009; EGFTEF.  
CC DR SMART: SM00181; EGF\_1.  
CC DR PROSITE: PS00022; EGF\_1; 1.  
CC DR PROSITE: PS01186; EGF\_2; 1.  
CC KW Signal; Growth factor; Heparin-binding; EGF-like domain;  
CC Transmembrane; Glycoprotein.  
CC FT SIGNAL 1 23  
CC FT PROPEP 24 62  
CC FT CHAIN 63 148  
CC FT PROPEP 149 208  
CC C-TERMINAL (POTENTIAL).  
C-TERMINAL (POTENTIAL).

```

FT DOMAIN 24 161 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 162 182 POTENTIAL.
FT DOMAIN 183 208 CYTOPLASMIC (POTENTIAL).
FT DOMAIN 104 144 EGF-LIKE.
FT DISULFID 108 121 BY SIMILARITY.
FT DISULFID 116 132 BY SIMILARITY.
FT DISULFID 134 143 BY SIMILARITY.
FT CARBOHYD 85 85 O-LINKED (GALNAC...) (BY SIMILARITY).
SQ SEQUENCE 208 AA; 22986 MW; 0A7DA97AE30C8967 CRC64;

Query Match 2.0%; Score 7; DB 1; Length 208;
Best Local Similarity 100.0%; Pred. No. 19;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 4 LPSVVL 10
| | | | |
Db 3 LPSVVL 9

RESULT 14
HBGF_RAT
ID HBGF_RAT STANDARD; PRT; 208 AA.
AC 006175;
DT 01-FEB-1995 (Rel. 31, Created)
DT 01-FEB-1995 (Rel. 31, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Heparin-binding EGF-like growth factor precursor (HB-EGF) (HBEGF).
GN DFR OR HBGF.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=SPRAGUE-DAWLEY; TISSUE=Macrophage;
RX MEDLINE=93135756; PubMed=7678488;
RA Abraham J.A., Damm D., Bejardi A., Miller J., Klagsbrun M.,
RA Ezekowitz R.A.B.;
RT Heparin-binding EGF-like growth factor: characterization of rat and
RT mouse cDNA clones, protein domain conservation across species, and
RT transcript expression in tissues";
RL Biochem. Biophys. Res. Commun. 190:125-133(1993).
CC -1- FUNCTION: MAY BE INVOLVED IN MACROPHAGE-MEDIATED CELLULAR
CC PROLIFERATION. IT IS MITOGENIC FOR FIBROBLASTS AND SMOOTH MUSCLE
CC BUT NOT ENDOTHELIAL CELLS. IT IS ABLE TO BIND EGF RECEPTORS WITH
CC HIGHER AFFINITY THAN EGF ITSELF AND IS A FAR MORE POTENT MITOGEN
CC FOR SMOOTH MUSCLE CELLS THAN EGF.
CC -1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN. MATURE HB-EGF IS
CC RELEASED INTO THE EXTRACELLULAR SPACE AND PROBABLY BINDS TO A
CC RECEPTOR.
CC -1- TISSUE SPECIFICITY: MOST ABUNDANT IN SKELETAL MUSCLE, LUNG, SPLEEN
CC BRAIN AND HEART.
CC -1- PTM: O-GLYCOSYLATED (BY SIMILARITY).
CC -1- SIMILARITY: CONTAINS 1 EGF-LIKE DOMAIN.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
CC EMBL: L05489; AAA81780.1; -.
CC PIR: JCI409; JCI409.
CC HSSP: Q99075; IADT.
CC InterPro: IPR000561; EGF-like.
CC Pfam: PF00008; EGF; 1.
CC SMART: SM00181; EGF; 1.
CC PROSITE: PS00022; EGF_1; 1.
CC PROSITE: PS01186; EGF_2; 1.
CC Signal; Growth factor; Heparin-binding; EGF-like domain;

```

```

KW Transmembrane; Glycoprotein.
FT SIGNAL 1 23 POTENTIAL.
FT PROPEP 24 62 BY SIMILARITY.
FT CHAIN 63 148 HEPARIN-BINDING EGF-LIKE GROWTH FACTOR.
FT PROPEP 149 208 C-TERMINAL (POTENTIAL).
FT DOMAIN 24 160 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 161 184 POTENTIAL.
FT DOMAIN 185 208 CYTOPLASMIC (POTENTIAL).
FT DOMAIN 104 144 EGF-LIKE.
FT DISULFID 108 121 BY SIMILARITY.
FT DISULFID 116 132 BY SIMILARITY.
FT DISULFID 134 143 BY SIMILARITY.
FT CARBOHYD 85 85 O-LINKED (GALNAC...) (BY SIMILARITY).
SQ SEQUENCE 208 AA; 22843 MW; DDB045E116D064C CRC64;

Query Match 2.0%; Score 7; DB 1; Length 208;
Best Local Similarity 100.0%; Pred. No. 19;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 4 LPSVVL 10
| | | | |
Db 3 LPSVVL 9

RESULT 15
HBX4_MOUSE
ID HBX4_MOUSE STANDARD; PRT; 250 AA.
AC P10284;
DT 01-MAR-1989 (Rel. 10, Created)
DT 01-MAR-1989 (Rel. 10, Last sequence update)
DT 01-OCT-1996 (Rel. 34, Last annotation update)
DE Homeobox protein Hox-B4 (Hox-2.6).
GN HOXB4 OR HOXB-4 OR HOX-2.6.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=89091992; PubMed=2463210;
RA Graham A., Papalopulu N., Lorimer J., McVey J.H., Tuddenham E.G.D.,
RA Krumlauf R.;
RT Characterization of a murine homeo box gene, Hox-2.6, related to the
RT Drosophila Deformed gene.";
RL Genes Dev. 2:1424-1438(1988).
CC -1- FUNCTION: SEQUENCE-SPECIFIC TRANSCRIPTION FACTOR WHICH IS PART OF
CC A DEVELOPMENTAL REGULATORY SYSTEM THAT PROVIDES CELLS WITH
CC SPECIFIC POSITIONAL IDENTITIES ON THE ANTERIOR-POSTERIOR AXIS.
CC -1- SUBCELLULAR LOCATION: Nuclear.
CC -1- SIMILARITY: BELONGS TO THE ANTP FAMILY OF HOMEBOX PROTEINS.
CC "DEFORMED" SUBFAMILY.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
CC EMBL: M36654; AAA37848.1; -.
CC PIR: A31757; A31757.
CC HSSP: P02833; 9ANT.
CC TRANSFAC: T01728; -.
CC MGD: MGI:96185; Hoxb4
CC InterPro: IPR001827; Antennapedia.
CC InterPro: IPR001356; Homeobox.
CC Pfam: PF00046; homeobox; 1.
CC PRINTS: PR00025; ANTENNAPEDIA.
CC PRINTS: PR00024; HOMEBOX.
CC SMART: SM00389; HOX; 1.
CC PROSITE: PS00027; HOMEBOX_1; 1.

```

DR PROSITE: PS00032; ANTENNAPEDIA; 1.  
 DR PROSITE: PS50071; HOMEBOX\_2; 1.  
 KW Homeobox; DNA-binding; Developmental protein; Nuclear protein;  
 KW Transcription regulation.  
 FT DOMAIN 15 138 PRO-RICH; PART OF THE TRANSCRIPTIONAL  
 FT DOMAIN 15 138 ACTIVATION DOMAIN.  
 FT DOMAIN 71 86 POLY-PRO.  
 FT DOMAIN 140 145 ANTP-TYPE HEXAPEPTIDE.  
 FT DNA BIND 161 220 HOMEBOX  
 SQ SEQUENCE 250 AA; 27519 MW; D09D477A0E585BE6 CRC64;

Query Match 2.0%; Score 7; DB 1; Length 250;  
 Best local Similarity 100.0%; Pred. No. 22;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 227 GGQRRES 233  
 Db 40 GGQRRES 46

Search completed: August 28, 2002, 11:27:18  
 Job time: 511 sec

... PAGE BLANK (USPTO)

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 28, 2002, 11:08:11 ; Search time 31 Seconds  
(without alignments)  
179.646 Million cell updates/sec

Title: US-09-853-526-70  
Perfect score: 1203  
Sequence: 1 MRLLPSVVLGTAPRYVLA.....NYLDAIVYVYEGKDDGG 228

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 231628 seqs, 24425594 residues

Total number of hits satisfying chosen parameters: 231628

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued\_Patents\_AA:\*  
1: /cgn2\_6/ptodata/2/1aa/5A.COMB.pep:\*  
2: /cgn2\_6/ptodata/2/1aa/5B.COMB.pep:\*  
3: /cgn2\_6/ptodata/2/1aa/6A.COMB.pep:\*  
4: /cgn2\_6/ptodata/2/1aa/6B.COMB.pep:\*  
5: /cgn2\_6/ptodata/2/1aa/PCTUS.COMB.pep:\*  
6: /cgn2\_6/ptodata/2/1aa/Dackfilest1.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1203	100.0	228	4	US-09-338-907-70 Sequence 70, Appl
2	1203	100.0	228	4	US-09-218-207-70 Sequence 70, Appl
3	1203	100.0	353	2	US-08-996-306-4 Sequence 4, Appl
4	1203	100.0	353	4	US-09-338-907-4 Sequence 4, Appl
5	1203	100.0	353	4	US-09-218-207-4 Sequence 4, Appl
6	1203	100.0	364	2	US-08-996-306-5 Sequence 5, Appl
7	1203	100.0	364	4	US-09-338-907-5 Sequence 5, Appl
8	1203	100.0	364	4	US-09-218-207-5 Sequence 5, Appl
9	1040	86.5	354	4	US-09-338-907-74 Sequence 74, Appl
10	1040	86.5	354	4	US-09-218-207-74 Sequence 74, Appl
11	976	81.1	185	4	US-09-338-907-136 Sequence 136, App
12	976	81.1	185	4	US-09-218-207-136 Sequence 136, App
13	976	81.1	315	4	US-09-338-907-134 Sequence 134, App
14	976	81.1	315	4	US-09-218-207-134 Sequence 134, App
15	970	80.6	300	4	US-09-338-907-135 Sequence 135, App
16	970	80.6	300	4	US-09-218-207-135 Sequence 135, App
17	830	69.0	291	4	US-09-338-907-127 Sequence 127, App
18	830	69.0	291	4	US-09-218-207-127 Sequence 127, App
19	820	68.2	182	4	US-09-338-907-133 Sequence 133, App
20	820	68.2	182	4	US-09-218-207-133 Sequence 133, App
21	657	54.6	261	4	US-09-338-907-128 Sequence 128, App
22	657	54.6	261	4	US-09-218-207-128 Sequence 128, App
23	597	49.6	238	4	US-09-338-907-126 Sequence 126, App
24	597	49.6	238	4	US-09-218-207-126 Sequence 126, App
25	449	37.3	97	4	US-09-338-907-132 Sequence 132, App
26	449	37.3	97	4	US-09-218-207-132 Sequence 132, App
27	332	27.6	77	4	US-09-338-907-125 Sequence 125, App

28	332	27.6	77	4	US-09-218-207-125	Sequence 125, App
29	330	27.4	66	4	US-09-338-907-131	Sequence 131, App
30	330	27.4	66	4	US-09-218-207-131	Sequence 131, App
31	329	27.3	68	4	US-09-338-907-130	Sequence 130, App
32	329	27.3	68	4	US-09-218-207-130	Sequence 130, App
33	329	27.3	90	4	US-09-338-907-129	Sequence 129, App
34	329	27.3	90	4	US-09-218-207-129	Sequence 129, App
35	269	22.4	374	2	US-08-454-267-2	Sequence 2, Appl
36	269	22.4	374	2	US-08-454-267-6	Sequence 6, Appl
37	269	22.4	374	2	US-08-941-319-2	Sequence 2, Appl
38	269	22.4	374	2	US-08-941-319-6	Sequence 6, Appl
39	269	22.4	374	4	US-09-035-098-2	Sequence 2, Appl
40	269	22.4	374	4	US-09-035-098-6	Sequence 6, Appl
41	269	22.4	374	4	US-09-215-252-5	Sequence 5, Appl
42	196	16.3	378	4	US-09-215-252-17	Sequence 17, Appl
43	187.5	15.6	376	4	US-09-215-252-13	Sequence 13, Appl
44	175.5	14.6	314	4	US-09-215-252-15	Sequence 15, Appl
45	171	14.2	295	2	US-08-454-267-7	Sequence 7, Appl

ALIGNMENTS

RESULT 1  
US-09-338-907-70  
; Sequence 70, Application US/09338907  
; Patent No. 6265546  
; GENERAL INFORMATION:  
; APPLICANT: Cohen, Daniel  
; APPLICANT: Blumenfeld, Marta  
; APPLICANT: Iliya, Chumakov  
; APPLICANT: Bouguenoret, Lydie  
; TITLE OF INVENTION: PROSTATE CANCER GENE  
; FILE REFERENCE: GENSET 18CPICP  
; CURRENT APPLICATION NUMBER: US/09/338, 907  
; CURRENT FILING DATE: 1999-06-23  
; EARLIER APPLICATION NUMBER: 08/996, 306  
; EARLIER FILING DATE: 1997-12-22  
; EARLIER APPLICATION NUMBER: 60/099, 658  
; EARLIER FILING DATE: 1998-09-09  
; EARLIER APPLICATION NUMBER: 09/218, 207  
; EARLIER FILING DATE: 1998-12-22  
; NUMBER OF SEQ ID NOS: 578  
; SOFTWARE: Patent.pm  
; SEQ ID NO 70  
; LENGTH: 228  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; US-09-338-907-70

Query Match	100.0%	Score 1203;	DB 4;	Length 228;
Best Local Similarity	100.0%	Pred. No. 2e-126;		
Matches 228;	Conservative	0;	Mismatches 0;	Indels 0;
			Gaps 0;	
OY	1	MRLLPSVVLGTAPRYVLA	NYLDAIVYVYEGKDDGG	60
DB	1	MRLLPSVVLGTAPRYVLA	NYLDAIVYVYEGKDDGG	60
OY	61	VOILLTGDPKKNENIYLANH	OSTVDWIVADILAI	RONALGHRVYLKEGKMLPLYGC 120
DB	61	VOILLTGDPKKNENIYLANH	OSTVDWIVADILAI	RONALGHRVYLKEGKMLPLYGC 120
OY	121	YFAOHGIVYKSAKNEKEM	RKKLOSYYDAGT	PMVLVIFPEBSTRNPEQTVLSASQAF 180
DB	61	VOILLTGDPKKNENIYLANH	OSTVDWIVADILAI	RONALGHRVYLKEGKMLPLYGC 120
OY	181	AAORGIAVLAHVLT	PRIKATHA	FDCKMNYLDAIVYVYEGKDDGG 228
DB	181	AAORGIAVLAHVLT	PRIKATHA	FDCKMNYLDAIVYVYEGKDDGG 228
RESULT	2			

US-09-218-207-70  
; Sequence 70, Application US/09218207  
; Patent No. 6346381  
; GENERAL INFORMATION:  
; APPLICANT: Cohen, Daniel  
; APPLICANT: Blumenfeld, Marta  
; APPLICANT: Ilya, Chumakov  
; APPLICANT: Bougueleret, Lydie  
; TITLE OF INVENTION: Prostate cancer gene  
; FILE REFERENCE: GENSET.018CPI  
; CURRENT APPLICATION NUMBER: US/09/218,207  
; EARLIER FILING DATE: 1998-12-22  
; EARLIER APPLICATION NUMBER: 60/996,306  
; EARLIER FILING DATE: 1997-12-22  
; EARLIER APPLICATION NUMBER: 60/099,658  
; NUMBER OF SEQ ID NOS: 578  
; SOFTWARE: Patent.pm  
; SEQ ID NO: 70  
; LENGTH: 228  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-218-207-70

Query Match 100.0%; Score 1203; DB 4; Length 228;  
Best Local Similarity 100.0%; Pred. No. 2e-126;  
Matches 228; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MRYLPSVVLGTAPYVVLAMGVRLLSAFLPAFYQALDDRLCYCVQSVLPEFEVYTG 60  
Db 1 MRYLPSVVLGTAPYVVLAMGVRLLSAFLPAFYQALDDRLCYCVQSVLPEFEVYTG 60  
QY 61 VOILLYGDLPRKNENIYILANHOSTVDIVADILAIQNALGHVRYLYLKGSLKWLPLYGC 120  
Db 61 VOILLYGDLPRKNENIYILANHOSTVDIVADILAIQNALGHVRYLYLKGSLKWLPLYGC 120  
QY 121 YFAOHGSIYVRSKAKFNEKEMRNKLOSYVDAGTPMYLYIFPEEGTRVPEOTKVLASQAF 180  
Db 121 YFAOHGSIYVRSKAKFNEKEMRNKLOSYVDAGTPMYLYIFPEEGTRVPEOTKVLASQAF 180  
QY 181 AAOAGLAVLKHVLTPIKATHVAFPDCKMKNYLDATYDVTYVYEGKDDGG 228  
Db 181 AAOAGLAVLKHVLTPIKATHVAFPDCKMKNYLDATYDVTYVYEGKDDGG 228

RESULT 3  
US-08-996-306-4  
; Sequence 4, Application US/08996306  
; Patent No. 5945522  
; GENERAL INFORMATION:  
; APPLICANT: Cohen, Daniel  
; APPLICANT: Chumakov, Ilya  
; APPLICANT: Blumenfeld, Marta  
; APPLICANT: Bougueleret, Lydie  
; TITLE OF INVENTION: Prostate cancer gene  
; NUMBER OF SEQUENCES: 68  
; CORRESPONDENCE ADDRESSES:  
; ADDRESSEE: Knobe, Martens, Olson & Bear  
; STREET: 501 West Broadway  
; CITY: San Diego  
; STATE: California  
; COUNTRY: USA  
; ZIP: 92101-3505  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; OPERATING SYSTEM: Win95  
; SOFTWARE: Word  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/996,306  
; FILING DATE:  
; CLASSIFICATION:

ATTORNEY/AGENT INFORMATION:  
; NAME: Israelsen, Ned A.  
; REGISTRATION NUMBER: 29,655  
; REFERENCE/DOCKET NUMBER: GENSET.018A  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (619) 235-8550  
; TELEFAX: (619) 235-0176  
; INFORMATION FOR SEQ ID NO: 4:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 353 amino acids  
; TYPE: AMINO ACID  
; STRANDEDNESS: SINGLE  
; TOPOLOGY: LINEAR  
; MOLECULE TYPE: PROTEIN  
; ORIGINAL SOURCE:  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: potential Transmembrane helix  
; LOCATION: 1..33  
; IDENTIFICATION METHOD: Rao and Argos method  
; FEATURE:  
; NAME/KEY: potential Transmembrane helix  
; LOCATION: 4..20  
; IDENTIFICATION METHOD: Klein, Kanehisa and Delisi method  
; FEATURE:  
; NAME/KEY: potential Transmembrane helix  
; LOCATION: 4..24  
; IDENTIFICATION METHOD: Eisenberg, Schwarz, Komaromy  
; IDENTIFICATION METHOD: and Wall method  
; FEATURE:  
; NAME/KEY: potential N-myristoylation site  
; LOCATION: 12  
; IDENTIFICATION METHOD: prosite match  
; FEATURE:  
; NAME/KEY: potential Transmembrane helix  
; LOCATION: 50..70  
; IDENTIFICATION METHOD: Eisenberg, Schwarz, Komaromy and Wall method  
; FEATURE:  
; NAME/KEY: potential N-glycosylation site  
; LOCATION: 57  
; IDENTIFICATION METHOD: prosite match  
; FEATURE:  
; NAME/KEY: potential Transmembrane helix  
; LOCATION: 76..96  
; IDENTIFICATION METHOD: Eisenberg, Schwarz, Komaromy and Wall method  
; FEATURE:  
; NAME/KEY: potential Tyrosine kinase phosphorylation site  
; LOCATION: 78  
; IDENTIFICATION METHOD: prosite match  
; FEATURE:  
; NAME/KEY: potential Casein kinase II phosphorylation site  
; LOCATION: 84  
; IDENTIFICATION METHOD: prosite match  
; FEATURE:  
; NAME/KEY: potential Leucine zipper pattern  
; LOCATION: 94..115  
; IDENTIFICATION METHOD: prosite match  
; FEATURE:  
; NAME/KEY: potential N-myristoylation site  
; LOCATION: 119  
; IDENTIFICATION METHOD: prosite match  
; FEATURE:  
; NAME/KEY: potential Protein kinase C phosphorylation site  
; LOCATION: 133  
; IDENTIFICATION METHOD: prosite match  
; FEATURE:  
; NAME/KEY: potential Casein kinase II phosphorylation site  
; LOCATION: 147  
; IDENTIFICATION METHOD: prosite match  
; FEATURE:  
; NAME/KEY: potential Protein kinase C phosphorylation site  
; LOCATION: 194  
; IDENTIFICATION METHOD: prosite match

```
FEATURE:
NAME/KEY: potential Tyrosine kinase phosphorylation site
LOCATION: 215
IDENTIFICATION METHOD: prosite match
FEATURE:
NAME/KEY: potential Tyrosine sulfatation site
LOCATION: 221
IDENTIFICATION METHOD: prosite match
FEATURE:
NAME/KEY: potential cAMP- and cGMP-dependent protein kinase phosphorylation site
LOCATION: 233
IDENTIFICATION METHOD: prosite match
FEATURE:
NAME/KEY: potential Casein kinase II phosphorylation site
LOCATION: 235
IDENTIFICATION METHOD: prosite match
FEATURE:
NAME/KEY: potential Protein kinase C phosphorylation site
LOCATION: 306
IDENTIFICATION METHOD: prosite match
FEATURE:
NAME/KEY: potential Transmembrane helix
LOCATION: 310..330
IDENTIFICATION METHOD: Eisenberg,Schwarz,Komarony and Wall method
FEATURE:
NAME/KEY: potential N-myristoylation site
LOCATION: 319
IDENTIFICATION METHOD: prosite match
FEATURE:
NAME/KEY: potential N-myristoylation site
LOCATION: 323
IDENTIFICATION METHOD: prosite match
FEATURE:
NAME/KEY: potential Amidation site
LOCATION: 329
IDENTIFICATION METHOD: prosite match
FEATURE:
NAME/KEY: potential Transmembrane helix
LOCATION: 333..353
IDENTIFICATION METHOD: Eisenberg,Schwarz,Komarony and Wall method
FEATURE:
NAME/KEY: potential N-myristoylation site
LOCATION: 341
IDENTIFICATION METHOD: prosite match
FEATURE:
NAME/KEY: potential Protein kinase C phosphorylation site
LOCATION: 350
IDENTIFICATION METHOD: prosite match
US-08-996-306-4
```

```
Query Match          100.0%; Score 1203; DB 2; Length 353;
Best Local Similarity 100.0%; Pred. No. 3,7e-126;
Matches 228; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY 1 MRYLLPSVVLGTAPTYVLAAGVWRLLSAFLPARFYQALDDRLCYCYOSMVLFFPENYNG 60
DB 1 MRYLLPSVVLGTAPTYVLAAGVWRLLSAFLPARFYQALDDRLCYCYOSMVLFFPENYNG 60
OY 61 VOILLYGDLPKKNENIYLANHOSYVDWIVADILAIROMALGHVRYVLEKGLKMLPLYGC 120
DB 61 VOILLYGDLPKKNENIYLANHOSYVDWIVADILAIROMALGHVRYVLEKGLKMLPLYGC 120
OY 121 YFAOHGGTYVRSKAFNEKEMRNKLOSYYDAGTPTMVLTFPEGSTRYNPEQTYVLSASQAF 180
DB 121 YFAOHGGTYVRSKAFNEKEMRNKLOSYYDAGTPTMVLTFPEGSTRYNPEQTYVLSASQAF 180
OY 181 AAORGLAVLKHYLTPRIKATHVAFDCMKNYLDAIDVTVVYEGKDDGG 228
DB 181 AAORGLAVLKHYLTPRIKATHVAFDCMKNYLDAIDVTVVYEGKDDGG 228
```

RESULT 4

```
US-09-338-907-4
; Sequence 4, Application US/09338907
; Patent No. 6265546
; GENERAL INFORMATION:
; APPLICANT: Cohen, Daniel
; APPLICANT: Blumenfeld, Marta
; APPLICANT: Ilya, Chumakov
; APPLICANT: Bougueleret, Lydie
; TITLE OF INVENTION: PROSTATE CANCER GENE
; FILE REFERENCE: GENSET.18CPLCP
; CURRENT APPLICATION NUMBER: US/09/338,907
; CURRENT FILING DATE: 1999-06-23
; EARLIER APPLICATION NUMBER: 08/996,306
; EARLIER FILING DATE: 1997-12-22
; EARLIER APPLICATION NUMBER: 60/099,658
; EARLIER FILING DATE: 1998-09-09
; EARLIER APPLICATION NUMBER: 09/218,207
; EARLIER FILING DATE: 1998-12-22
; NUMBER OF SEQ ID NOS: 578
; SOFTWARE: Patent.pm
; SEQ ID NO 4
; LENGTH: 353
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: HELIX
; LOCATION: 1..33
; OTHER INFORMATION: Rao and Argos identification method, potential helix
; FEATURE:
; NAME/KEY: HELIX
; LOCATION: 4..20
; OTHER INFORMATION: Klein, Kanehisa and Delisi identification method, potent
; FEATURE:
; NAME/KEY: HELIX
; LOCATION: 4..24
; OTHER INFORMATION: Eisenberg, Schwarz, Komarony, Wall identification method,
; FEATURE:
; NAME/KEY: MYRISTATE
; LOCATION: 12..16
; OTHER INFORMATION: Prosite match
; FEATURE:
; NAME/KEY: HELIX
; LOCATION: 50..70
; OTHER INFORMATION: Eisenberg, Schwarz, Komarony, Wall identification method,
; FEATURE:
; NAME/KEY: CARBOHYD
; LOCATION: 57..59
; OTHER INFORMATION: Prosite match
; FEATURE:
; NAME/KEY: HELIX
; LOCATION: 76..96
; OTHER INFORMATION: Eisenberg, Schwarz, Komarony, Wall identification method,
; FEATURE:
; NAME/KEY: PHOSPHORYLATION
; LOCATION: 78
; OTHER INFORMATION: potential Tyrosine kinase site, Prosite match
; FEATURE:
; NAME/KEY: PHOSPHORYLATION
; LOCATION: 84
; OTHER INFORMATION: potential Tyrosine kinase site, Prosite match
; FEATURE:
; NAME/KEY: SITE
; LOCATION: 94..115
; OTHER INFORMATION: potential leucine zipper site, Prosite match
; FEATURE:
; NAME/KEY: MYRISTATE
; LOCATION: 119..123
; OTHER INFORMATION: potential site, Prosite match
; FEATURE:
; NAME/KEY: PHOSPHORYLATION
; LOCATION: 133
; OTHER INFORMATION: potential protein kinase C, Prosite match
; FEATURE:
```

```
NAME/KEY: PHOSPHORYLATION
LOCATION: 147
OTHER INFORMATION: potential caseine kinase II site, Prosite match
FEATURE:
NAME/KEY: PHOSPHORYLATION
LOCATION: 194
OTHER INFORMATION: potential protein kinase C, Prosite match
FEATURE:
NAME/KEY: PHOSPHORYLATION
LOCATION: 215
OTHER INFORMATION: potential Tyrosine kinase site, Prosite match
FEATURE:
NAME/KEY: SUFATATION
LOCATION: 221
OTHER INFORMATION: Prosite match
FEATURE:
NAME/KEY: PHOSPHORYLATION
LOCATION: 233
OTHER INFORMATION: potential cAMP and cGMP dependant protein kinase site, Prosite match
FEATURE:
NAME/KEY: PHOSPHORYLATION
LOCATION: 235
OTHER INFORMATION: potential caseine kinase II site, Prosite match
FEATURE:
NAME/KEY: PHOSPHORYLATION
LOCATION: 306
OTHER INFORMATION: potential protein kinase C, Prosite match
FEATURE:
NAME/KEY: HELIX
LOCATION: 310..330
OTHER INFORMATION: Eisenberg, Schwarz, Komarony, Wall Identification method, Prosite match
FEATURE:
NAME/KEY: MYRISTATE
LOCATION: 319..323
OTHER INFORMATION: Prosite match
FEATURE:
NAME/KEY: MYRISTATE
LOCATION: 323..327
OTHER INFORMATION: Prosite match
FEATURE:
NAME/KEY: AMIDATION
LOCATION: 329
OTHER INFORMATION: Prosite match
FEATURE:
NAME/KEY: HELIX
LOCATION: 333..353
OTHER INFORMATION: Eisenberg, Schwarz, Komarony, Wall Identification method, Prosite match
FEATURE:
NAME/KEY: MYRISTATE
LOCATION: 341..345
OTHER INFORMATION: Prosite match
FEATURE:
NAME/KEY: PHOSPHORYLATION
LOCATION: 350
OTHER INFORMATION: potential protein kinase C, Prosite match
US-09-338-907-4
```

Query Match 100.0%; Score 1203; DB 4; Length 353;  
Best Local Similarity 100.0%; Pred. No. 3.7e-126;  
Matches 228; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY 1 MRYLLPSVVLGTAPYVLAAGVWRLSAFLPARFYQALDDRLCYVQSWLFFFEVNTG 60
D 1 MRYLLPSVVLGTAPYVLAAGVWRLSAFLPARFYQALDDRLCYVQSWLFFFEVNTG 60
QY 61 VQILLVGDLPKKNENITYLANHSTVDIADILAIROMALGHVRYVLKESLWLPPLYGC 120
D 61 VQILLVGDLPKKNENITYLANHSTVDIADILAIROMALGHVRYVLKESLWLPPLYGC 120
QY 121 YFAOHGSIYVRSKAKFEKEMKRNKLOSVDAGTPTMYLVIPPEGRYRPEQTKVLSASQAF 180
D 121 YFAOHGSIYVRSKAKFEKEMKRNKLOSVDAGTPTMYLVIPPEGRYRPEQTKVLSASQAF 180
```

```
QY 181 AAGRLAVLKHVLPRIKATHVAEDCKMKNYLDIYDVTVYVEGKDDG 228
D 181 AAGRLAVLKHVLPRIKATHVAEDCKMKNYLDIYDVTVYVEGKDDG 228

RESULT 5
US-09-218-207-4
Sequence 4, Application US/09218207
Patent No. 6346381
GENERAL INFORMATION:
APPLICANT: Cohen, Daniel
APPLICANT: Blumenfeld, Marta
APPLICANT: Ilya, Chumakov
APPLICANT: Bougueleret, Lydie
TITLE OF INVENTION: Prostate cancer gene
FILE REFERENCE: GENSET 018CPI
CURRENT APPLICATION NUMBER: US/09/218, 207
CURRENT FILING DATE: 1998-12-22
EARLIER APPLICATION NUMBER: 08/996, 306
EARLIER FILING DATE: 1997-12-22
EARLIER APPLICATION NUMBER: 60/099, 658
EARLIER FILING DATE: 1998-09-09
NUMBER OF SEQ ID NOS: 578
SOFTWARE: PatentL.pm
SEQ ID NO 4
LENGTH: 353
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: HELIX
LOCATION: 1..33
OTHER INFORMATION: Rao and Argos Identification method, potential helix
FEATURE:
NAME/KEY: HELIX
LOCATION: 4..20
OTHER INFORMATION: Klein, Kanehisa and Delisi Identification method, potential helix
FEATURE:
NAME/KEY: MYRISTATE
LOCATION: 12..16
OTHER INFORMATION: Prosite match
FEATURE:
NAME/KEY: HELIX
LOCATION: 50..70
OTHER INFORMATION: Eisenberg, Schwarz, Komarony, Wall Identification method, Prosite match
FEATURE:
NAME/KEY: CARBOHYD
LOCATION: 57..59
OTHER INFORMATION: Prosite match
FEATURE:
NAME/KEY: HELIX
LOCATION: 76..96
OTHER INFORMATION: Eisenberg, Schwarz, Komarony, Wall Identification method, Prosite match
FEATURE:
NAME/KEY: PHOSPHORYLATION
LOCATION: 78
OTHER INFORMATION: potential Tyrosine kinase site, Prosite match
FEATURE:
NAME/KEY: PHOSPHORYLATION
LOCATION: 84
OTHER INFORMATION: potential caseine kinase II site, Prosite match
FEATURE:
NAME/KEY: SITE
LOCATION: 94..115
OTHER INFORMATION: potential Leucine zipper site, Prosite match
FEATURE:
NAME/KEY: MYRISTATE
LOCATION: 119..123
OTHER INFORMATION: potential site, Prosite match
```



FEATURE: NAME/KEY: PHOSPHORYLATION  
LOCATION: 133  
OTHER INFORMATION: potential protein kinase C, Prosite match  
FEATURE: NAME/KEY: PHOSPHORYLATION  
LOCATION: 147  
OTHER INFORMATION: potential caseine kinase II site, Prosite match  
FEATURE: NAME/KEY: PHOSPHORYLATION  
LOCATION: 194  
OTHER INFORMATION: potential protein kinase C, Prosite match  
FEATURE: NAME/KEY: PHOSPHORYLATION  
LOCATION: 215  
OTHER INFORMATION: potential tyrosine kinase site, Prosite match  
FEATURE: NAME/KEY: SULFATATION  
LOCATION: 221  
OTHER INFORMATION: Prosite match  
FEATURE: NAME/KEY: PHOSPHORYLATION  
LOCATION: 233  
OTHER INFORMATION: potential CAMP and cGMP dependant protein kinase site, Prosite match  
FEATURE: NAME/KEY: PHOSPHORYLATION  
LOCATION: 235  
OTHER INFORMATION: potential caseine kinase II site, Prosite match  
FEATURE: NAME/KEY: PHOSPHORYLATION  
LOCATION: 306  
OTHER INFORMATION: potential protein kinase C, Prosite match  
FEATURE: NAME/KEY: HELIX  
LOCATION: 310..330  
OTHER INFORMATION: Eisenberg, Schwarz, Komaromy, Wall identification method, F  
FEATURE: NAME/KEY: MYRISTATE  
LOCATION: 319..323  
OTHER INFORMATION: Prosite match  
FEATURE: NAME/KEY: MYRISTATE  
LOCATION: 323..327  
OTHER INFORMATION: Prosite match  
FEATURE: NAME/KEY: AMIDATION  
LOCATION: 329  
OTHER INFORMATION: Prosite match  
FEATURE: NAME/KEY: HELIX  
LOCATION: 333..353  
OTHER INFORMATION: Eisenberg, Schwarz, Komaromy, Wall identification method, F  
FEATURE: NAME/KEY: MYRISTATE  
LOCATION: 341..345  
OTHER INFORMATION: Prosite match  
FEATURE: NAME/KEY: PHOSPHORYLATION  
LOCATION: 350  
OTHER INFORMATION: potential protein kinase C, Prosite match  
US-09-218-207-4

Query Match 100.0%; Score 1203; DB 4; Length 353;  
Best local Similarity 100.0%; Pred. No. 3.7e-126;  
Matches 228; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MRLPLSVVLGTAPRYVLAAGVWRLLSAFLPARFYQALDDRLCYVQSMVLEFFENYTG 60  
DB 1 MRLPLSVVLGTAPRYVLAAGVWRLLSAFLPARFYQALDDRLCYVQSMVLEFFENYTG 60  
QY 61 VOILLXGDLPRKNENIYLANHOSVDWIVADILAIRONALGHVRYVLEKGLKMLPLYGC 120  
DB 61 VOILLXGDLPRKNENIYLANHOSVDWIVADILAIRONALGHVRYVLEKGLKMLPLYGC 120

DB 61 VOILLXGDLPRKNENIYLANHOSVDWIVADILAIRONALGHVRYVLEKGLKMLPLYGC 120  
QY 121 YFAOHGIGYVKKRSKAKFNEMKRLQSYVDAGTAPMLVIFPESTRNPEOTKVLASQAF 180  
DB 121 YFAOHGIGYVKKRSKAKFNEMKRLQSYVDAGTAPMLVIFPESTRNPEOTKVLASQAF 180  
QY 181 AAOGRGLAVLKHVLTFRKATHVAFDCMKNYLDAIVDTVYVEGKDDGG 228  
DB 181 AAOGRGLAVLKHVLTFRKATHVAFDCMKNYLDAIVDTVYVEGKDDGG 228

RESULT 6  
US-08-996-306-5  
Sequence 5, Application US/08996306  
Patent No. 5945522  
GENERAL INFORMATION:  
APPLICANT: Cohen, Daniel  
APPLICANT: Chumakov, Ilya  
APPLICANT: Blumenfeld, Marla  
APPLICANT: Bougueleret, Lydie  
TITLE OF INVENTION: Prostate cancer gene  
NUMBER OF SEQUENCES: 68  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Knobbe, Martens, Olson & Bear  
STREET: 501 West Broadway  
CITY: San Diego  
STATE: California  
COUNTRY: USA  
ZIP: 92101-3505  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy Disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: Win95  
SOFTWARE: Word  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/996,306  
FILING DATE:  
CLASSIFICATION:  
ATTORNEY/AGENT INFORMATION:  
NAME: Israelien, Ned A.  
REGISTRATION NUMBER: 29,655  
REFERENCE/DOCKET NUMBER: GENSET.018A  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (619) 235-8550  
TELEFAX: (619) 235-0176  
INFORMATION FOR SEQ ID NO: 5:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 364 amino acids  
TYPE: AMINO ACID  
STRANDEDNESS: SINGLE  
TOPOLOGY: LINEAR  
MOLECULE TYPE: PROTEIN  
ORIGINAL SOURCE:  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: potential protein  
LOCATION: 1..364  
US-08-996-306-5

Query Match 100.0%; Score 1203; DB 2; Length 364;  
Best local Similarity 100.0%; Pred. No. 3.9e-126;  
Matches 228; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MRLPLSVVLGTAPRYVLAAGVWRLLSAFLPARFYQALDDRLCYVQSMVLEFFENYTG 60  
DB 12 MRLPLSVVLGTAPRYVLAAGVWRLLSAFLPARFYQALDDRLCYVQSMVLEFFENYTG 71  
QY 61 VOILLXGDLPRKNENIYLANHOSVDWIVADILAIRONALGHVRYVLEKGLKMLPLYGC 120  
DB 72 VOILLXGDLPRKNENIYLANHOSVDWIVADILAIRONALGHVRYVLEKGLKMLPLYGC 131  
QY 121 YFAOHGIGYVKKRSKAKFNEMKRLQSYVDAGTAPMLVIFPESTRNPEOTKVLASQAF 180

```
Db 132 YFAOHGIIYKRSKAKFNEKEKRNKLOSVDAGTPMYLVIFPEGTRYNPEOTKVLASQAF 191
Qy 181 AAORGLAVLKHVLTLPRIKATHVAFDCMKNYLDAIYDVTVVYEGKDDG 228
Db 192 AAORGLAVLKHVLTLPRIKATHVAFDCMKNYLDAIYDVTVVYEGKDDG 239
```

```
RESULT 7
US-09-338-907-5
; Sequence 5, Application US/09338907
; Patent No. 6265546
; GENERAL INFORMATION:
; APPLICANT: Cohen, Daniel
; APPLICANT: Blumenfeld, Marta
; APPLICANT: Ilya, Chumakov
; APPLICANT: Bougueleret, Lydie
; TITLE OF INVENTION: PROSTATE CANCER GENE
; FILE REFERENCE: GENSET.18CP1CP
; CURRENT APPLICATION NUMBER: US/09/338,907
; EARLIER FILING DATE: 1999-06-23
; EARLIER APPLICATION NUMBER: 08/996,306
; EARLIER FILING DATE: 1997-12-22
; EARLIER APPLICATION NUMBER: 60/099,658
; EARLIER FILING DATE: 1998-09-09
; EARLIER APPLICATION NUMBER: 09/218,207
; NUMBER OF SEQ ID NOS: 578
; SOFTWARE: Patent.pm
; SEQ ID NO 5
; LENGTH: 364
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-338-907-5
```

```
Query Match 100.0%; Score 1203; DB 4; Length 364;
Best Local Similarity 100.0%; Pred. No. 3.9e-126;
Matches 228; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MRYLLPSVVLGTAPTYVLAAGVWRLSAPLPARFYQALDDRLCYCYQSWLFFPENYTG 60
Db 12 MRYLLPSVVLGTAPTYVLAAGVWRLSAPLPARFYQALDDRLCYCYQSWLFFPENYTG 71
Qy 61 VOILLYGDLPKKNENIYYLANHOSVDWIVADILAIROMALGHVRYVLEKGLKWLPLYGC 120
Db 72 VOILLYGDLPKKNENIYYLANHOSVDWIVADILAIROMALGHVRYVLEKGLKWLPLYGC 131
Qy 121 YFAOHGIIYKRSKAKFNEKEKRNKLOSVDAGTPMYLVIFPEGTRYNPEOTKVLASQAF 180
Db 132 YFAOHGIIYKRSKAKFNEKEKRNKLOSVDAGTPMYLVIFPEGTRYNPEOTKVLASQAF 191
Qy 181 AAORGLAVLKHVLTLPRIKATHVAFDCMKNYLDAIYDVTVVYEGKDDG 228
Db 192 AAORGLAVLKHVLTLPRIKATHVAFDCMKNYLDAIYDVTVVYEGKDDG 239
```

```
RESULT 8
US-09-218-207-5
; Sequence 5, Application US/09218207
; Patent No. 6346381
; GENERAL INFORMATION:
; APPLICANT: Cohen, Daniel
; APPLICANT: Blumenfeld, Marta
; APPLICANT: Ilya, Chumakov
; APPLICANT: Bougueleret, Lydie
; TITLE OF INVENTION: Prostate cancer gene
; FILE REFERENCE: GENSET.018CP1
; CURRENT APPLICATION NUMBER: US/09/218,207
; EARLIER FILING DATE: 1998-12-22
; EARLIER APPLICATION NUMBER: 08/996,306
; EARLIER FILING DATE: 1997-12-22
; EARLIER APPLICATION NUMBER: 60/099,658
```

```
; EARLIER FILING DATE: 1998-09-09
; NUMBER OF SEQ ID NOS: 578
; SOFTWARE: Patent.pm
; SEQ ID NO 5
; LENGTH: 364
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-218-207-5
```

```
Query Match 100.0%; Score 1203; DB 4; Length 364;
Best Local Similarity 100.0%; Pred. No. 3.9e-126;
Matches 228; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MRYLLPSVVLGTAPTYVLAAGVWRLSAPLPARFYQALDDRLCYCYQSWLFFPENYTG 60
Db 12 MRYLLPSVVLGTAPTYVLAAGVWRLSAPLPARFYQALDDRLCYCYQSWLFFPENYTG 71
Qy 61 VOILLYGDLPKKNENIYYLANHOSVDWIVADILAIROMALGHVRYVLEKGLKWLPLYGC 120
Db 72 VOILLYGDLPKKNENIYYLANHOSVDWIVADILAIROMALGHVRYVLEKGLKWLPLYGC 131
Qy 121 YFAOHGIIYKRSKAKFNEKEKRNKLOSVDAGTPMYLVIFPEGTRYNPEOTKVLASQAF 180
Db 132 YFAOHGIIYKRSKAKFNEKEKRNKLOSVDAGTPMYLVIFPEGTRYNPEOTKVLASQAF 191
Qy 181 AAORGLAVLKHVLTLPRIKATHVAFDCMKNYLDAIYDVTVVYEGKDDG 228
Db 192 AAORGLAVLKHVLTLPRIKATHVAFDCMKNYLDAIYDVTVVYEGKDDG 239
```

```
RESULT 9
US-09-338-907-74
; Sequence 74, Application US/09338907
; Patent No. 6265546
; GENERAL INFORMATION:
; APPLICANT: Cohen, Daniel
; APPLICANT: Blumenfeld, Marta
; APPLICANT: Ilya, Chumakov
; APPLICANT: Bougueleret, Lydie
; TITLE OF INVENTION: PROSTATE CANCER GENE
; FILE REFERENCE: GENSET.18CP1CP
; CURRENT APPLICATION NUMBER: US/09/338,907
; EARLIER FILING DATE: 1999-06-23
; EARLIER APPLICATION NUMBER: 08/996,306
; EARLIER FILING DATE: 1997-12-22
; EARLIER APPLICATION NUMBER: 60/099,658
; EARLIER FILING DATE: 1998-09-09
; EARLIER APPLICATION NUMBER: 09/218,207
; NUMBER OF SEQ ID NOS: 578
; SOFTWARE: Patent.pm
; SEQ ID NO 74
; LENGTH: 354
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-338-907-74
```

```
Query Match 86.5%; Score 1040; DB 4; Length 354;
Best Local Similarity 85.9%; Pred. No. 6e-108;
Matches 195; Conservative 19; Mismatches 13; Indels 0; Gaps 0;

Qy 1 MRYLLPSVVLGTAPTYVLAAGVWRLSAPLPARFYQALDDRLCYCYQSWLFFPENYTG 60
Db 1 MRYLLPSVVLGTAPTYVLAAGVWRLSAPLPARFYQALDDRLCYCYQSWLFFPENYTG 60
Qy 61 VOILLYGDLPKKNENIYYLANHOSVDWIVADILAIROMALGHVRYVLEKGLKWLPLYGC 120
Db 61 VOILLYGDLPKKNENIYYLANHOSVDWIVADILAIROMALGHVRYVLEKGLKWLPLYGC 120
Qy 121 YFAOHGIIYKRSKAKFNEKEKRNKLOSVDAGTPMYLVIFPEGTRYNPEOTKVLASQAF 180
```

```
Db      121 YFAOHGIIYVKRSKAKFNDEKMSKLOSYYNAGTPLYIFPESTRYNATYTKLLSASQAF 180
QY      181 AAORGLAVLKHVLTPIRIKATHVAFDCMKMYLDAIYDVTVYVEGKDDG 227
Db      181 AAORGLAVLKHVLTPIRIKATHVAFDSMKSHLDIYDVTVYVEGNEKG 227
```

```
RESULT 10
US-09-218-207-74
; Sequence 74, Application US/09218207
; Patent No. 6346381
; GENERAL INFORMATION:
; APPLICANT: Cohen, Daniel
; APPLICANT: Blumenfeld, Marta
; APPLICANT: Ilya, Chumakov
; APPLICANT: Bougueleret, Lydie
; TITLE OF INVENTION: Prostate cancer gene
; FILE REFERENCE: GENSET.018CPI
; CURRENT APPLICATION NUMBER: US/09/218,207
; CURRENT FILING DATE: 1998-12-22
; EARLIER APPLICATION NUMBER: 08/996,306
; EARLIER FILING DATE: 1997-12-22
; EARLIER APPLICATION NUMBER: 60/099,658
; EARLIER FILING DATE: 1998-09-09
; NUMBER OF SEQ ID NOS: 578
; SOFTWARE: Patent.pm
; SEQ ID NO 74
; LENGTH: 354
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-218-207-74
```

```
Query Match      86.5%; Score 1040; DB 4; Length 354;
Best Local Similarity 85.9%; Pred. No. 6e-108;
Matches 195; Conservative 19; Mismatches 13; Indels 0; Gaps 0;
```

```
QY      1 MRYLLPSVVLGTAPTYVLAAGVWRLSAPLPARFYQALDDRLCYQOSMVLFFFEENYTG 60
Db      1 MRYLLPSVVLGSAPTYVLAAGVWRLSAPLPARFYQALDDRLCYQOSMVLFFFEENYTG 60
QY      61 VOILLYGDLPKKNENIYLANHOSYVDWIVADILAIRONALGHVRYVLEKGLKWLPLYGC 120
Db      61 VOILLYGDLPKKNENIYLANHOSYVDWIVADILAIRONALGHVRYVLEKGLKWLPLYGC 120
QY      121 YFAOHGIIYVKRSKAKFNDEKMSKLOSYYNAGTPLYIFPESTRYNATYTKLLSASQAF 180
Db      121 YFAOHGIIYVKRSKAKFNDEKMSKLOSYYNAGTPLYIFPESTRYNATYTKLLSASQAF 180
QY      181 AAORGLAVLKHVLTPIRIKATHVAFDCMKMYLDAIYDVTVYVEGKDDG 227
Db      181 AAORGLAVLKHVLTPIRIKATHVAFDSMKSHLDIYDVTVYVEGNEKG 227
```

```
RESULT 11
US-09-338-907-136
; Sequence 136, Application US/09338907
; Patent No. 6265546
; GENERAL INFORMATION:
; APPLICANT: Cohen, Daniel
; APPLICANT: Blumenfeld, Marta
; APPLICANT: Ilya, Chumakov
; APPLICANT: Bougueleret, Lydie
; TITLE OF INVENTION: PROSTATE CANCER GENE
; FILE REFERENCE: GENSET.18CPICP
; CURRENT APPLICATION NUMBER: US/09/338,907
; CURRENT FILING DATE: 1999-06-23
; EARLIER APPLICATION NUMBER: 08/996,306
; EARLIER FILING DATE: 1997-12-22
; EARLIER APPLICATION NUMBER: 60/099,658
; EARLIER FILING DATE: 1998-09-09
; EARLIER APPLICATION NUMBER: 09/218,207
; EARLIER FILING DATE: 1998-12-22
```

```
; NUMBER OF SEQ ID NOS: 578
; SOFTWARE: Patent.pm
; SEQ ID NO 136
; LENGTH: 185
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: 81..83
; OTHER INFORMATION: Box I
; FEATURE:
; NAME/KEY: SITE
; LOCATION: 160..165
; OTHER INFORMATION: Box II
US-09-338-907-136
```

```
Query Match      81.1%; Score 976; DB 4; Length 185;
Best Local Similarity 100.0%; Pred. No. 3.3e-101;
Matches 185; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      1 MRYLLPSVVLGTAPTYVLAAGVWRLSAPLPARFYQALDDRLCYQOSMVLFFFEENYTG 60
Db      1 MRYLLPSVVLGTAPTYVLAAGVWRLSAPLPARFYQALDDRLCYQOSMVLFFFEENYTG 60
QY      61 VOILLYGDLPKKNENIYLANHOSYVDWIVADILAIRONALGHVRYVLEKGLKWLPLYGC 120
Db      61 VOILLYGDLPKKNENIYLANHOSYVDWIVADILAIRONALGHVRYVLEKGLKWLPLYGC 120
QY      121 YFAOHGIIYVKRSKAKFNDEKMSKLOSYYNAGTPLYIFPESTRYNATYTKLLSASQAF 180
Db      121 YFAOHGIIYVKRSKAKFNDEKMSKLOSYYNAGTPLYIFPESTRYNATYTKLLSASQAF 180
QY      181 AAORG 185
Db      181 AAORG 185
```

```
RESULT 12
US-09-218-207-136
; Sequence 136, Application US/09218207
; Patent No. 6346381
; GENERAL INFORMATION:
; APPLICANT: Cohen, Daniel
; APPLICANT: Blumenfeld, Marta
; APPLICANT: Ilya, Chumakov
; APPLICANT: Bougueleret, Lydie
; TITLE OF INVENTION: Prostate cancer gene
; FILE REFERENCE: GENSET.018CPI
; CURRENT APPLICATION NUMBER: US/09/218,207
; CURRENT FILING DATE: 1998-12-22
; EARLIER APPLICATION NUMBER: 08/996,306
; EARLIER FILING DATE: 1997-12-22
; EARLIER APPLICATION NUMBER: 60/099,658
; EARLIER FILING DATE: 1998-09-09
; NUMBER OF SEQ ID NOS: 578
; SOFTWARE: Patent.pm
; SEQ ID NO 136
; LENGTH: 185
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: 81..83
; OTHER INFORMATION: Box I
; FEATURE:
; NAME/KEY: SITE
; LOCATION: 160..165
; OTHER INFORMATION: Box II
US-09-218-207-136
```

```
Query Match      81.1%; Score 976; DB 4; Length 185;
```



; EARLIER FILING DATE: 1997-12-22  
; EARLIER APPLICATION NUMBER: 60/099, 658  
; EARLIER FILING DATE: 1998-09-09  
; EARLIER APPLICATION NUMBER: 09/218, 207  
; EARLIER FILING DATE: 1998-12-22  
; NUMBER OF SEQ ID NOS: 578  
; SOFTWARE: Patent.pm  
; SEQ ID NO 135  
; LENGTH: 300  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: SITE  
; LOCATION: 81..83  
; OTHER INFORMATION: Box 1  
; FEATURE:  
; NAME/KEY: SITE  
; LOCATION: 160..165  
; OTHER INFORMATION: Box 11  
US-09-338-907-135

Query Match 80.6%; Score 970; DB 4; Length 300;  
Best Local Similarity 100.0%; Pred. No. 3.1e-100;  
Matches 184; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
OY 1 MRYLLPSVVLIGTAPFYVLAWGVWRLSAPFLPARYQALDDRLXCVQSMVLEFFENTG 60  
Db 1 MRYLLPSVVLIGTAPFYVLAWGVWRLSAPFLPARYQALDDRLXCVQSMVLEFFENTG 60  
OY 61 VOILLYGDLPRKNENIYLANHOSVDWIVADILAIRONALGHVRYVLEKGIKMLPLYGC 120  
Db 61 VOILLYGDLPRKNENIYLANHOSVDWIVADILAIRONALGHVRYVLEKGIKMLPLYGC 120  
OY 121 YFAOHGGIYVRSKAFENKRNKLOSTYVDAGTPLYVIFPEGTRYNPQOTKVLASQAF 180  
Db 121 YFAOHGGIYVRSKAFENKRNKLOSTYVDAGTPLYVIFPEGTRYNPQOTKVLASQAF 180  
OY 181 AAQR 184  
Db 181 AAQR 184

Search completed: August 28, 2002, 11:08:12  
Job time: 291 sec

**THIS PAGE BLANK (USPTO)**

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 28, 2002, 11:07:34 ; Search time 75.55 Seconds  
(without alignments)  
335.206 Million cell updates/sec

Title: US-09-853-526-70

Perfect score: 1203

Sequence: 1 MRYLPSVVLGTAPRYVLA.....NYLDALIVYVYEGKDDG 228

Scoring table:

BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 747574 seqs, 11073796 residues

Total number of hits satisfying chosen parameters: 747574

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database :

A.Geneseq\_032802:\*  
1: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1980.DAT:\*  
2: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1981.DAT:\*  
3: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1982.DAT:\*  
4: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1983.DAT:\*  
5: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1984.DAT:\*  
6: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1985.DAT:\*  
7: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1986.DAT:\*  
8: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1987.DAT:\*  
9: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1988.DAT:\*  
10: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1989.DAT:\*  
11: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1990.DAT:\*  
12: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1991.DAT:\*  
13: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1992.DAT:\*  
14: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1993.DAT:\*  
15: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1994.DAT:\*  
16: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1995.DAT:\*  
17: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1996.DAT:\*  
18: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1997.DAT:\*  
19: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1998.DAT:\*  
20: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA2000.DAT:\*  
21: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA2001.DAT:\*  
22: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA2001.DAT:\*

Prod. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	1203	100.0	228	20	AAV36740	Truncated PGI prot
2	1203	100.0	352	20	AAV36728	Human PGI protein
3	1203	100.0	353	22	AAV36591	Human polypeptide
4	1203	100.0	353	22	AAV36532	Human secreted pro
5	1203	100.0	353	20	AAV36729	Human PGI protein
6	1203	100.0	364	22	AAV36729	Human acyltransfer
7	1203	100.0	364	22	AAV36729	Human lysophosphat
8	1203	100.0	364	22	AAV36729	Human protein sequ
9	1200	99.8	353	22	AAV36595	Human polypeptide
10	1192	99.1	372	22	AAV36741	Mouse PGI protein
11	1040	86.5	354	20	AAV36741	

12	976	81.1	185	20	AAV36753	PGI splice variant
13	976	81.1	315	20	AAV36751	PGI splice variant
14	970	80.6	300	20	AAV36752	PGI splice variant
15	830	69.0	291	20	AAV36744	PGI splice variant
16	820	68.2	182	20	AAV36750	PGI splice variant
17	759	63.1	269	22	AAE09653	Human gene 15 enco
18	657	54.6	261	20	AAV36745	PGI splice variant
19	597	49.6	238	20	AAV36743	PGI splice variant
20	449	37.3	96	20	AAV36749	PGI splice variant
21	387	32.2	257	22	ABG01360	Novel human diapo
22	387	32.2	980	22	ABG11977	Novel human diapo
23	332	27.6	77	20	AAV36742	PGI splice variant
24	330	27.4	66	20	AAV36748	PGI splice variant
25	329	27.3	68	20	AAV36747	PGI splice variant
26	329	27.3	90	20	AAV36746	PGI splice variant
27	312	25.9	378	21	AAV36740	PGI splice variant
28	307	25.5	352	21	AAV36742	PGI splice variant
29	304	25.3	341	21	AAV36742	PGI splice variant
30	291	24.2	375	21	AAV36734	PGI splice variant
31	291	24.2	375	21	AAV36742	PGI splice variant
32	288	23.9	351	21	AAV36735	PGI splice variant
33	288	23.9	375	21	AAV36742	PGI splice variant
34	285	23.7	351	21	AAV36742	PGI splice variant
35	275	22.9	343	21	AAV36736	PGI splice variant
36	275	22.9	343	21	AAV36743	PGI splice variant
37	269	22.4	374	22	AAV36742	PGI splice variant
38	267	22.2	374	15	AAV36742	PGI splice variant
39	264	21.9	374	21	AAV36742	PGI splice variant
40	253.5	21.1	294	21	AAV36742	PGI splice variant
41	252	20.9	374	21	AAV36742	PGI splice variant
42	242.5	20.2	377	21	AAV36742	PGI splice variant
43	242	20.1	377	17	AAV36742	PGI splice variant
44	239.5	19.9	387	21	AAV36742	PGI splice variant
45	238.5	19.8	389	21	AAV36742	PGI splice variant

ALIGNMENTS

RESULT 1	
AAV36740	standard; Protein; 228 AA.
AAV36740;	
AC	27-SEP-1999 (first entry)
DT	Truncated PGI protein sequence.
DE	
KW	PGI gene; diallelic marker; human; PSA; PGI-related diallelic marker;
KW	cancer; prostate cancer; diagnosis; therapy; prostate specific antigen.
OS	Homo sapiens.
XX	
XX	
PN	W09932644-A2.
PD	01-JUL-1999.
PF	22-DEC-1998; 98WO-1B02133.
PR	09-SEP-1998; 98US-0099658.
PR	22-DEC-1997; 97US-0996306.
PA	(GEST ) GENSET.
PI	Blumenfeld M, Bougueleret L, Chumakov I, Cohen D;
PI	WPI: 1999-405178/34.
DR	N-PSDB; AA00926.
DR	
PT	Use of a prostate cancer associated gene and diallelic markers
PT	derived from it
XX	

PS Claim 7; Page 208; 385pp; English.  
 XX  
 CC The invention relates to a mammalian PGI gene and protein, and a set of  
 CC PGI biallelic markers. The PGI polynucleotide and biallelic markers are  
 CC used in a hybridisation assay, a sequencing assay, or in an  
 CC allele-specific amplification assay for determining the identity of a  
 CC nucleotide at a PGI-related biallelic marker. The methods can be used to  
 CC detect and to assess the risk of developing cancer or prostate cancer.  
 CC Early-stage diagnosis of prostate cancer relies on prostate specific  
 CC antigen (PSA) dosage. However, the effectiveness of this is limited due  
 CC to its inability to discriminate between malignant and non-malignant  
 CC effects of the organ. A need exists for both a reliable diagnostic  
 CC procedure which would enable early-stage diagnosis, and for preventative  
 CC and curative treatments of the disease. The PGI gene can be used for  
 CC detection of prostate cancer, and the risk of developing it in the  
 CC future, and can also be used to determine therapies for the disease.  
 CC  
 XX  
 SQ Sequence 228 AA;

Query Match 100.0%; Score 1203; DB 20; Length 228;  
 Best Local Similarity 100.0%; Pred. No. 2.5e-124;  
 Matches 228; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MRYLLPSVVLGTRPTVYVLAMGWRLLSAFLPARFYQALDRLYCYQSMVLEFFENYTG 60  
 DB 1 MRYLLPSVVLGTRPTVYVLAMGWRLLSAFLPARFYQALDRLYCYQSMVLEFFENYTG 60  
 QY 61 VOILLYDLDPRKNENIITYLANHOSTVDMIVADILAIRONALGHVRYLKEGLKWLPLYGC 120  
 DB 61 VOILLYDLDPRKNENIITYLANHOSTVDMIVADILAIRONALGHVRYLKEGLKWLPLYGC 120  
 QY 121 YFAOHGITYVRSKAFNEKEMRNKLSQSYVDAGTPMYLVIFPEEGRYNPEOTKYLSASQAF 180  
 DB 121 YFAOHGITYVRSKAFNEKEMRNKLSQSYVDAGTPMYLVIFPEEGRYNPEOTKYLSASQAF 180  
 QY 121 YFQHGIIYVRSKAFNEKEMRNKLSQSYVDAGTPMYLVIFPEEGRYNPEOTKYLSASQAF 180  
 DB 121 YFQHGIIYVRSKAFNEKEMRNKLSQSYVDAGTPMYLVIFPEEGRYNPEOTKYLSASQAF 180  
 QY 181 AAGRGVLAVLKHVLPRIKATHVAFDCKKNYLDIYDVTVVYEGKDDGG 228  
 DB 181 AAGRGVLAVLKHVLPRIKATHVAFDCKKNYLDIYDVTVVYEGKDDGG 228

## RESULT 2

AAY36728 2  
 ID AAY36728 standard; Protein; 352 AA.

AC AAY36728;  
 XX  
 DT 27-SEP-1999 (first entry)  
 XX  
 DE Human PGI protein sequence.  
 XX  
 KW PGI gene; biallelic marker; human; PSA; PGI-related biallelic marker;  
 KW cancer; prostate cancer; diagnosis; therapy; prostate specific antigen.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO932644-A2.  
 XX  
 PD 01-JUL-1999.  
 XX  
 PF 22-DEC-1998; 98WO-IB02133.  
 XX  
 PR 09-SEP-1998; 98US-0099658.  
 PR 22-DEC-1997; 97US-0096306.  
 XX  
 PA (GEST ) GENSET.  
 XX  
 PI Blumenfeld M, Bougueleret L, Chumakov I, Cohen D;  
 XX  
 DR WPI: 1999-405178/34.  
 DR N-PSDB; AA200872.  
 XX  
 PT Use of a prostate cancer associated gene and biallelic markers

PT derived from it  
 XX  
 XX Claim 7; Page 190; 385pp; English.  
 XX  
 CC The invention relates to a mammalian PGI gene and protein, and a set of  
 CC PGI biallelic markers. The PGI polynucleotide and biallelic markers are  
 CC used in a hybridisation assay, a sequencing assay, or in an  
 CC allele-specific amplification assay for determining the identity of a  
 CC nucleotide at a PGI-related biallelic marker. The methods can be used to  
 CC detect and to assess the risk of developing cancer or prostate cancer.  
 CC Early-stage diagnosis of prostate cancer relies on prostate specific  
 CC antigen (PSA) dosage. However, the effectiveness of this is limited due  
 CC to its inability to discriminate between malignant and non-malignant  
 CC effects of the organ. A need exists for both a reliable diagnostic  
 CC procedure which would enable early-stage diagnosis, and for preventative  
 CC and curative treatments of the disease. The PGI gene can be used for  
 CC detection of prostate cancer, and the risk of developing it in the  
 CC future, and can also be used to determine therapies for the disease.  
 CC  
 XX  
 SQ Sequence 352 AA;

Query Match 100.0%; Score 1203; DB 20; Length 352;  
 Best Local Similarity 100.0%; Pred. No. 4.6e-124;  
 Matches 228; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MRYLLPSVVLGTRPTVYVLAMGWRLLSAFLPARFYQALDRLYCYQSMVLEFFENYTG 60  
 DB 1 MRYLLPSVVLGTRPTVYVLAMGWRLLSAFLPARFYQALDRLYCYQSMVLEFFENYTG 60  
 QY 61 VOILLYDLDPRKNENIITYLANHOSTVDMIVADILAIRONALGHVRYLKEGLKWLPLYGC 120  
 DB 61 VOILLYDLDPRKNENIITYLANHOSTVDMIVADILAIRONALGHVRYLKEGLKWLPLYGC 120  
 QY 121 YFAOHGITYVRSKAFNEKEMRNKLSQSYVDAGTPMYLVIFPEEGRYNPEOTKYLSASQAF 180  
 DB 121 YFAOHGITYVRSKAFNEKEMRNKLSQSYVDAGTPMYLVIFPEEGRYNPEOTKYLSASQAF 180  
 QY 121 YFQHGIIYVRSKAFNEKEMRNKLSQSYVDAGTPMYLVIFPEEGRYNPEOTKYLSASQAF 180  
 DB 121 YFQHGIIYVRSKAFNEKEMRNKLSQSYVDAGTPMYLVIFPEEGRYNPEOTKYLSASQAF 180  
 QY 181 AAGRGVLAVLKHVLPRIKATHVAFDCKKNYLDIYDVTVVYEGKDDGG 228  
 DB 181 AAGRGVLAVLKHVLPRIKATHVAFDCKKNYLDIYDVTVVYEGKDDGG 228

## RESULT 3

AAM39591 3  
 ID AAM39591 standard; Protein; 353 AA.

AC AAM39591;  
 XX  
 DT 22-OCT-2001 (first entry)  
 XX  
 DE Human polypeptide SEQ ID NO 2736.  
 XX  
 KW Human; nootropic; immunosuppressant; cytostatic; gene therapy; cancer;  
 KW peripheral nervous system; neuropathy; central nervous system; CNS;  
 KW Alzheimer's; Parkinson's disease; Huntington's disease; hemostatic;  
 KW amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;  
 KW chemokinetic; thrombolytic; drug screening; arthritis; inflammation;  
 KW Leukemia.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO20015312-A1.  
 XX  
 PD 26-JUL-2001.  
 XX  
 PF 26-DEC-2000; 2000WO-US34263.  
 XX  
 PR 21-JAN-2000; 2000US-0488725.  
 PR 25-APR-2000; 2000US-0552317.  
 PR 09-JUL-2000; 2000US-0598042.  
 PR 19-JUL-2000; 2000US-0620312.  
 PR 03-AUG-2000; 2000US-0653450.



14-SEP-2000; 2000US-0662191.  
PR 19-OCT-2000; 2000US-0693036.  
PR 29-NOV-2000; 2000US-0727344.  
XX  
XX  
XX (HSE-) HSEQ INC.  
XX Tang YF, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;  
PI Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J;  
PI Zhao QA, Zhou P, Goodrich R, Drmanac RT;  
XX  
XX WPI: 2001-442253/47.  
DR N-PSDB: AA158747.  
XX  
XX Novel nucleic acids and polypeptides, useful for treating disorders  
PT such as central nervous system injuries -  
PS Example 4: SEQ ID NO 2736; 10078pp: English.  
XX  
XX The invention relates to human nucleic acids (AA157798-AA161369) and  
CC the encoded polypeptides (AA136642-AA142213) with nootropic,  
CC immunosuppressant and cytostatic activity. The polynucleotides are useful  
CC in gene therapy. A composition containing a polypeptide or polynucleotide  
CC of the invention may be used to treat diseases of the peripheral nervous  
CC system, such as peripheral nervous injuries, peripheral neuropathy and  
CC localised neuropathies and central nervous system diseases, such as  
CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic  
CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the  
CC utilisation of the activities such as: immune system suppression,  
CC Activin/Inhibin activity, chemotactic/chemokinetic activity, haemostatic  
CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,  
CC assays for receptor activity, arthritis and inflammation, leukaemias and  
CC C.N.S disorders.  
CC Note: The sequence data for this patent did not form part of the printed  
CC specification.  
XX  
XX Sequence 353 AA:

Query Match 100.0%; Score 1203; DB 22; Length 353;  
Best Local Similarity 100.0%; Pred. No. 4.6e-124;  
Matches 228; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MRLTSPVVLGTAPTYVLAAGWRLLSAFLPARFQALDDRLCYQSWVLFEPENYTG 60  
DB 1 mryllpsvvlgtaptyvlawgwrlsalfparfygalddrlcyqswvlfefenytg 60  
QY 61 VOILLGDLPRKNENITVLANHOSVDWIVADILAIRONALGHVRYVLEGLKWLPLYGC 120  
DB 61 vqillgdlprknenitvylanhsqstvdwlvadilairqnalghvryvllkeglkwlplygc 120  
QY 121 YFAOHGIGIVKRSKFNEMKRLQSYVDAGTPMVLTFPESTRNPEQTKVLSAQAF 180  
DB 121 yfaohgigivkrsakfnemrkllqsyvdagtpmvltfpestrnpeqtkvlsasqaf 180  
QY 181 AAQRGLAVLKHVLTPIKATHVAFDCMKNYLDIYDVTVVYVYEGSKDDGG 228  
DB 181 aagrglavlkhvltprikathvafdcmknyldaiydvttvvyegskddgg 228

RESULT 4  
AAB85532  
ID AAB85532 standard; protein; 353 AA.  
XX  
XX AAB85532;  
AC  
XX 25-SEP-2001 (first entry)  
XX  
XX Human secreted protein (clone id HLTRO84).  
XX  
XX Secreted protein; immunosuppressive; antirheumatic; antineoplastic;  
KW antiproliferative; cytostatic; cardiac; vasotonic; cerebroprotective;  
KW nootropic; neuroprotective; antibacterial; virocidic; fungicide; human;  
KW ophthalmological; gene therapy.

XX Homo sapiens.  
OS  
XX WO200155430-A1.  
PN  
XX  
XX 02-AUG-2001.  
PD  
XX  
XX 17-JAN-2001; 2001WO-US01431.  
PF  
XX  
XX 31-JAN-2000; 2000US-0179065.  
PR 04-FEB-2000; 2000US-0180628.  
PR 12-SEP-2000; 2000US-0231968.  
XX  
XX (HUMA-) HUMAN GENOME SCI INC.  
PA  
XX Rosen CA, Komatsoulis GA, Baker KP, Birse CE, Soppet DR, Olsen HS;  
PI Moore PA, Wei P, Edner R, Duan DR, Shi Y, Choi GH, Fiscella M;  
PI Ni J, Ruben SM, Barash SC;  
XX  
XX WPI: 2001-476220/51.  
DR N-PSDB: AAH46942.

17 isolated nucleic acid molecules encoding human secreted proteins,  
used to preventing, treating or ameliorating a medical condition -  
Claim 11: Page 450-451; 482pp; English.

The invention provides novel human secreted proteins and polynucleotides  
encoding them. The secreted proteins can be expressed by standard  
recombinant methodology. The secreted proteins and polynucleotides are  
used to prevent, treat or ameliorate a medical condition in e.g. humans,  
mice, rabbits, goats, horses, cats, dogs, chickens or sheep. They can  
also be used in diagnosing a pathological condition. The antibodies to  
the proteins can also be used in alleviating symptoms associated with the  
disorders and in diagnostic immunoassays e.g. radioimmunoassays or enzyme  
linked immunosorbent assays (ELISA). Disorders which are diagnosed or  
treated include autoimmune diseases e.g. rheumatoid arthritis,  
hyperproliferative disorders e.g. neoplasms of the breast or liver,  
cardiovascular disorders e.g. cardiac arrest, cerebrovascular disorders  
e.g. cerebral ischemia, angiodenesis, nervous system disorders e.g.  
Alzheimer's disease, infections caused by bacteria, viruses and fungi and  
ocular disorders e.g. corneal infection. The polypeptides can also be  
used to aid wound healing and epithelial cell proliferation, to prevent  
skin aging due to sunburn, to maintain organs before transplantation, for  
supporting cell culture of primary tissues, to regenerate tissues and in  
chemotaxis. The polypeptides can also be used as a food additive or  
preservative to increase or decrease storage capabilities. The present  
sequence represents a human secreted protein.

Query Match 100.0%; Score 1203; DB 22; Length 353;  
Best Local Similarity 100.0%; Pred. No. 4.6e-124;  
Matches 228; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MRLTSPVVLGTAPTYVLAAGWRLLSAFLPARFQALDDRLCYQSWVLFEPENYTG 60  
DB 1 mryllpsvvlgtaptyvlawgwrlsalfparfygalddrlcyqswvlfefenytg 60  
QY 61 VOILLGDLPRKNENITVLANHOSVDWIVADILAIRONALGHVRYVLEGLKWLPLYGC 120  
DB 61 vqillgdlprknenitvylanhsqstvdwlvadilairqnalghvryvllkeglkwlplygc 120  
QY 121 YFAOHGIGIVKRSKFNEMKRLQSYVDAGTPMVLTFPESTRNPEQTKVLSAQAF 180  
DB 121 yfaohgigivkrsakfnemrkllqsyvdagtpmvltfpestrnpeqtkvlsasqaf 180  
QY 181 AAQRGLAVLKHVLTPIKATHVAFDCMKNYLDIYDVTVVYVYEGSKDDGG 228  
DB 181 aagrglavlkhvltprikathvafdcmknyldaiydvttvvyegskddgg 228

```

RESULT      5
AAB85552
ID    AAB85552 standard; protein; 353 AA.
XX
XX
AC    AAB85552;
XX
XX
DT    25-SEP-2001 (first entry)
XX
XX
DE    Human secreted protein (clone Id HSL1A81).
XX
XX
KW    Secreted protein; immunosuppressive; antiarthritic; antirheumatic;
KW    antiproliferative; cytostatic; cadiant; vasotropic; cerebroprotective;
XX    neotropic; neuroprotective; antibacterial; virucide; fungicide; human;
KW    ophthalmological; gene therapy.
XX
XX
OS    Homo sapiens.
XX
XX
PN    MO200155430-A1.
XX
XX
PD    02-AUG-2001.
XX
XX
PE    17-JAN-2001; 2001MO-US01431.
XX
XX
PR    31-JAN-2000; 2000US-0179065.
PR    04-FEB-2000; 2000US-0180628.
PR    12-SEP-2000; 2000US-0231968.
XX
XX
PA    (HUMA-) HUMAN GENOME SCI INC.
XX
XX
PI    Rosen CA, Komatsoulis GA, Baker KP, Birse CE, Soppet DR, Olsen HS;
PI    Moore PA, Wei P, Ebner R, Duan DR, Shi Y, Choi GH, Fiscella M,
PI    Ni J, Ruben SM, Barash SC;
XX
XX
DR    WPI: 2001-476220/51.
DR    N-PSDB; AAH46962.
XX
XX
PT    17 isolated nucleic acid molecules encoding human secreted proteins,
PT    used to preventing, treating or ameliorating a medical condition -
XX
XX
PS    Claim 11; Page 465-466; 483pp; English.
XX
XX
CC    The invention provides novel human secreted proteins and polynucleotides
CC    encoding them. The secreted proteins can be expressed by standard
CC    recombinant methodology. The secreted proteins and polynucleotides are
CC    used to prevent, treat or ameliorate a medical condition in e.g. humans,
CC    mice, rabbits, goats, horses, cats, dogs, chickens or sheep. They can
CC    also be used in diagnosing a pathological condition. The antibodies to
CC    the proteins can be used in alleviating symptoms associated with the
CC    disorders and in diagnostic immunoassays e.g. radioimmunoassays or enzyme
CC    linked immunosorbent assays (ELISA). Disorders which are diagnosed or
CC    treated include autoimmune diseases e.g. rheumatoid arthritis,
CC    hyperproliferative disorders e.g. neoplasms of the breast or liver,
CC    cardiovascular disorders e.g. cardiac arrest, cerebrovascular disorders
CC    e.g. cerebral ischemia, angioneuosis, nervous system disorders e.g.
CC    Alzheimer's disease, infections caused by bacteria, viruses and fungi and
CC    ocular disorders e.g. corneal infection. The polypeptides can also be
CC    used to aid wound healing and epithelial cell proliferation, to prevent
CC    skin aging due to sunburn, to maintain organs before transplantation, for
CC    supporting cell culture of primary tissues, to regenerate tissues and in
CC    chemotaxis. The polypeptides can also be used as a food additive or
CC    preservative to increase or decrease storage capabilities. The present
CC    sequence represents a human secreted protein.
XX
XX
Sequence    353 AA;

```

```

OY 61 VOIILXGDLPRKKENIITYLANHOSVDMVYADIIILRQALGHVRYVYKESGKMLPIYGC 120
Db 61 VqIIILYgdlpRkKenIIyLanHqSVdMvYAdIIILRqALghVryVYkEgSkMlPlYgc 120
OY 121 YFAOHGCIYVRSARFENEKEMRNKLOSVDYADCTPMVLVYFPEGSTRYNPEQTVLSASQAF 180
Db 121 YfaOhgciYvRsArFEnEkEmrNkLoSVdYAdCTpMvLVyFpEGsTrYnPeQTVlSasqAf 180
OY 181 AAORGLAVIKHVLTPRIKATHVAFDCMKNYLDAIVDYTVYVYSGKDDG 228
Db 181 AaOrGLaVikHvLTpRIkaThVaFdcmkNyLdaIVdyTVyVYsGkDdG 228

RESULT 6
AAV36729
ID AAV36729 standard; Protein; 364 AA.
XX
AC AAV36729;
XX
DT 27-SEP-1999 (first entry)
XX
DE Human PGI protein sequence.
XX
KW PGI gene; biallelic marker; human; PSA; PGI-related biallelic marker;
KW cancer; prostate cancer; diagnosis; therapy; prostate specific antigen.
OS Homo sapiens.
XX
PN MO9932644-A2.
XX
PD 01-JUL-1999.
XX
PF 22-DEC-1998; 98WO-IB02133.
PR 09-SEP-1998; 98US-0099658.
PR 22-DEC-1997; 97US-0096306.
XX
PA (GEST ) GENSET.
PI Blumenfeld M, Bougueleret L, Chumakov I, Cohen D;
XX
DR WPI; 1999-405178/34.
DR N-PSDB; AAZ00870.
XX
PT Use of a prostate cancer associated gene and biallelic markers
PT derived from it
XX
PS Claim 7; Page 190-191; 385pp; English.
XX
CC The invention relates to a mammalian PGI gene and protein, and a set of
CC PGI biallelic markers. The PGI polynucleotide and biallelic markers are
CC used in a hybridisation assay, a sequencing assay, or in an
CC allele-specific amplification assay for determining the identity of a
CC nucleotide at a PGI-related biallelic marker. The methods can be used to
CC detect and to assess the risk of developing cancer or prostate cancer.
CC Early-stage diagnosis of prostate cancer relies on prostate specific
CC antigen (PSA) dosage. However, the effectiveness of this is limited due
CC to its inability to discriminate between malignant and non-malignant
CC affections of the organ. A need exists for both a reliable diagnostic
CC procedure which would enable early-stage diagnosis, and for preventative
CC and curative treatments of the disease. The PGI gene can be used for
CC detection of prostate cancer, and the risk of developing it in the
CC future, and can also be used to determine therapies for the disease.
XX
SQ Sequence 364 AA;

```

```

Db      12 myllpsvlllgtpcylawgvwrlisafiparfygaddrlcyvgsmviffenyg 71
Oy      61 VOILLYGLDLPKKNKENTITLANHQSIVDMIVADILAIRONALGHVRYVLEKGLKWLPLYGC 120
Db      72 vqlllygdipkknkentiylanbqstcvdwiadilaiprnaighvryvlekgikwlpilygc 131
Oy      121 YFAOHGIIYVKRSKAFNEKEMRNKLQSYVDAGTPMYLVIFPEGSTRYNPEQTVLSASQAF 180
Db      132 yfagnggilyvkrksakfnekenmklsqsyvdagtpmylvifpegtlrynpeqtkvlsasqaf 191
Oy      181 AAQGLAVLKHVLPRIKATHVAFDCMKNYLDIYDVTVVYEGKDDG 228
Db      192 aagrglawlkhvlprikathvafdcmknyldiaydvtrvvyegkddgg 239

RESULT 7
AAM50128
ID      AAM50128 standard; Protein: 364 AA.
XX
AC      AAM50128;
XX
DT      21-DEC-2001 (first entry)
XX
DE      Human acyltransferase 27417.
XX
KW      Acyltransferase 27417; human; cell proliferation; cell migration;
KW      cell differentiation; ovary cancer; brain cancer; colon cancer;
KW      lung cancer; tumour; metastasis; sarcoma; carcinoma;
KW      adenocarcinoma; antitumour; lipid metabolism; diagnosis; therapy.
XX
OS      Homo sapiens.
XX

Key      Location/Qualifiers
FT      Peptide      1..49
FT      /label= Signal_peptide
FT      Protein      50..364
FT      /label= Mature_protein
FT      Domain      50..320
FT      /label= N-terminal_domain
FT      Domain      321..337
FT      /label= Transmembrane_domain
FT      Domain      338..364
FT      /label= C-terminal_domain
FT      Domain      71..363
FT      /label= Acyltransferase_domain
FT      Modified-site 68..71
FT      /note= "Asn is N-glycosylated"
FT      Modified-site 241..244
FT      /note= "O-phosphorylated by cAMP- and
FT      cGMP-dependent protein kinase"
FT      Modified-site 11..13
FT      /note= "O-phosphorylated by protein kinase C"
FT      Modified-site 144..146
FT      /note= "O-phosphorylated by protein kinase C"
FT      Modified-site 205..207
FT      /note= "O-phosphorylated by protein kinase C"
FT      Modified-site 317..319
FT      /note= "O-phosphorylated by protein kinase C"
FT      Modified-site 361..363
FT      /note= "O-phosphorylated by protein kinase C"
FT      Modified-site 95..98
FT      /note= "O-phosphorylated by casein kinase II"
FT      Modified-site 158..161
FT      /note= "O-phosphorylated by casein kinase II"
FT      Modified-site 246..249
FT      /note= "O-phosphorylated by casein kinase II"
FT      Modified-site 82..89
FT      /note= "O-phosphorylated by tyrosine kinase"
FT      Modified-site 219..226
FT      /note= "O-phosphorylated by tyrosine kinase"
FT      Modified-site 23..28
FT      /note= "N-myristoylated"

```

```

FT      Modified-site 130..135
FT      /note= "N-myristoylated"
FT      Modified-site 330..335
FT      /note= "N-myristoylated"
FT      Modified-site 352..357
FT      /note= "N-myristoylated"
FT      Modified-site 230..343
FT      /note= "Amidation"
FT      Peptide      105..126
FT      /note= "predicted leucine zipper"
XX
XX      WO200173051-A2.
XX
XX      PD      04-OCT-2001.
XX
XX      PE      26-MAR-2001; 2001WO-US09633.
XX
XX      PR      24-MAR-2000; 2000US-192092P.
XX
XX      PA      (MILL-) MILLENNIUM PHARM INC.
XX
XX      PI      Meyers RA, Rudolph-Owen L, Macbeth KJ, Tsai F, Williamson M;
XX
XX      DR      WPI, 2001-626267/72.
XX
XX      DR      N-PSDB; AAH27094.
XX
XX      PS      Claim 9; Fig 8A-B; 150pp; English.
XX
XX      CC      The present sequence is that of a novel human protein, termed
XX      CC      27417, which shows the structural characteristics of members of the
XX      CC      acyltransferase family, including an acyltransferase domain. The
XX      CC      invention provides novel acyltransferase 27417 nucleic acids and
XX      CC      polypeptides, as well as methods for detecting their presence, and
XX      CC      methods for screening for compounds that modulate their expression
XX      CC      or activity. Such compounds can be used to treat conditions
XX      CC      related to aberrant activity or expression of 27417 protein or
XX      CC      nucleic acid, such as conditions involving deficient cellular
XX      CC      proliferation, migration and/or differentiation. Also provided are
XX      CC      methods for inhibiting the proliferation or migration, or inducing
XX      CC      the killing, of a 27417-expressing cell, e.g. a hyperproliferative
XX      CC      and/or metastatic cell, by contact with a compound that modulates
XX      CC      the activity or expression of the 27417 protein or nucleic acid.
XX      CC      The 27417-expressing cell is found in the healthy or diseased
XX      CC      heart, blood vessels, kidney, skeletal muscle, brain or liver, or
XX      CC      especially in a solid tumour, a soft tissue tumour or a metastatic
XX      CC      lesion, a sarcoma, a carcinoma or an adenocarcinoma, and in
XX      CC      particular in a hyperproliferative and/or metastatic cell found in
XX      CC      ovarian, brain, colon or lung cancer. The compounds may also be
XX      CC      used to modulate lipid metabolism in a 27417-expressing cell.
XX      CC      Methods of disease diagnosis, e.g. by determining the
XX      CC      presence of a genetic alteration in a 27417 polypeptide, and for
XX      CC      evaluating the efficacy of a treatment of a disorder, are also
XX      CC      provided.
XX
XX      SQ      Sequence      364 AA:

Query Match      100.0%; Score 1203; DB 22; Length 364;
Best Local Similarity 100.0%; Pred. No. 4,86-124;
Matches 228; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Oy      1 MRYLLPSVLLGTAPRYVAMGVWRLLSAFLPAPFOALDDRLRYCYOSMVEFFENYTG 60
Db      12 myllpsvlllgtpcylawgvwrlisafiparfygaddrlcyvgsmviffenyg 71
Oy      61 VOILLYGLDLPKKNKENTITLANHQSIVDMIVADILAIRONALGHVRYVLEKGLKWLPLYGC 120
Db      72 vqlllygdipkknkentiylanbqstcvdwiadilaiprnaighvryvlekgikwlpilygc 131

```

QY 121 YFAOHGCIYVKSARENEKEMRNKLSQSYVDAGTPMWLVIFPEGTRYNPEQTKVLSASQAF 180  
 |||  
 DB 132 yfaqhggilyvkrsakfnekemrnlqgyvdagcpmylvifpegcrlypeqtkvlaasqaf 191  
 |||  
 QY 181 AAGRGALVLTHTPRKATHTVAFDDCKNLDATYDVTYVYEGKDDCG 228  
 |||  
 DB 192 aagrglavltkhvltprkathvaIdcmknyldalylvlvvgyekddgg 239  
 |||  
 RESULT 8  
 AAU00665  
 ID AAU00665 standard; Protein: 364 AA.  
 XX  
 AC AAU00665;  
 XX  
 DT 07-SEP-2001 (first entry)  
 XX  
 DE Human lysophosphatidic acid acyltransferase isoform LPAAT-epsilon.  
 XX  
 KM Lysophosphatidic acid acyltransferase epsilon; LPAAT-epsilon; LPA; PA;  
 KM lysophosphatidic acid; phosphatidic acid; acylation; cellular activation;  
 KM phospholipid signalling; mitogenesis; inflammation; autoimmune disease;  
 KM oncology; cancer; obesity; gene therapy.  
 XX  
 OS Homo sapiens.  
 XX  
 PN W0200134782-A1.  
 XX  
 PD 17-MAY-2001.  
 XX  
 PE 02-NOV-2000; 2000MO-US30193.  
 XX  
 PR 09-NOV-1999; 99US-0436919.  
 XX  
 PA (CELL-) CELL THERAPEUTICS INC.  
 XX  
 PI Leung DW;  
 XX  
 DR WPI: 2001-335920/35.  
 DR N-PSDB; AAS00649.  
 XX  
 PT Novel isolated human isoform of lysophosphatidic acid  
 PT acyltransferase-epsilon useful for diagnostic, therapeutic and  
 PT screening purposes -  
 XX  
 PS Claim 1; Fig 1; 48pp; English.  
 XX  
 CC The sequence represents a human lysophosphatidic acid acyltransferase  
 CC (LPAAT) isoform, LPAAT-epsilon. LPAAT catalyses the acylation of  
 CC lysophosphatidic acid (LPA) to phosphatidic acid (PA). LPA and PA have  
 CC been identified as phospholipid signalling molecules that affect a wide  
 CC range of biological responses. PA is involved in cellular activation and  
 CC mitogenesis. Compounds that block PA generation and hence diminish lipid  
 CC biosynthesis and the signal involved in cell activation are of  
 CC therapeutic interest in the areas of inflammation and oncology (e.g.  
 CC autoimmune diseases and cancer) as well as obesity treatment.  
 CC LPAAT-epsilon and its corresponding DNA can be used in screening assays  
 CC to detect agents that stimulate or inhibit the activity of LPAAT and,  
 CC therefore, PA. The DNA is useful in tests to detect the presence or  
 CC expression of LPAAT-epsilon in relation to certain diseases and  
 CC conditions, and in disease prevention and treatment. The sequences of the  
 CC invention are also useful for diagnosis of diseases and conditions in  
 CC which the expression of LPAAT enzyme is abnormal.  
 CC  
 SQ Sequence 364 AA;  
 Query Match 100.0%; Score 1203; DB 22; Length 364;  
 Best local Similarity 100.0%; Pred. No. 4,8e-124;  
 Matches 228; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 MRYLPSVVLGTAPTYVLAMGVRLISAFLPARYOALDDRLCYQVSMVLFPEFNYTG 60

DB 12 myrlpsvvlgtaptyvlamgvrrllsatlparfygaldrlycvysmvlffenytg 71  
 |||  
 QY 61 VOILLYGDLPRKKNENITYLANHSTVDMIVADILAIQNALGHRVYKLGKMLPYGC 120  
 |||  
 DB 72 vqlllygdldpknkenillylanhstvdmivadilaitqnalghrvyikgkmlpygc 131  
 |||  
 QY 121 YFAOHGCIYVKSARENEKEMRNKLSQSYVDAGTPMWLVIFPEGTRYNPEQTKVLSASQAF 180  
 |||  
 DB 132 yfaqhggilyvkrsakfnekemrnlqgyvdagcpmylvifpegcrlypeqtkvlaasqaf 191  
 |||  
 QY 181 AAGRGALVLTHTPRKATHTVAFDDCKNLDATYDVTYVYEGKDDCG 228  
 |||  
 DB 192 aagrglavltkhvltprkathvaIdcmknyldalylvlvvgyekddgg 239  
 |||  
 RESULT 9  
 AAB93595  
 ID AAB93595 standard; Protein: 353 AA.  
 XX  
 AC AAB93595;  
 XX  
 DT 26-JUN-2001 (first entry)  
 XX  
 DE Human protein sequence SEQ ID NO:13028.  
 XX  
 KM Human; primer; detection; diagnosis; antisense therapy; gene therapy.  
 XX  
 OS Homo sapiens.  
 XX  
 PN EP1074617-A2.  
 XX  
 PD 07-FEB-2001.  
 XX  
 PE 28-JUL-2000; 2000EP-0116126.  
 XX  
 PR 29-JUL-1999; 99JP-0248036.  
 PR 27-AUG-1999; 99UP-0300253.  
 PR 11-JAN-2000; 2000JP-0118776.  
 PR 02-MAY-2000; 2000JP-0183767.  
 PR 09-JUN-2000; 2000JP-0241899.  
 XX  
 PA (HELI-) HELIX RES INST.  
 XX  
 PI Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;  
 PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;  
 XX  
 DR WPI: 2001-318749/34.  
 XX  
 CC Primer sets for synthesizing polynucleotides, particularly the 5602  
 CC full-length cDNAs defined in the specification, and for the detection  
 CC and/or diagnosis of the abnormality of the proteins encoded by the  
 CC full-length cDNAs -  
 XX  
 PS Claim 8; SEQ ID 13028; 2537pp + CD ROM; English.  
 XX  
 CC The present invention describes primer sets for synthesizing 5602  
 CC full-length cDNAs defined in the specification. Where a primer set  
 CC comprises: (a) an oligo-dT primer and an oligonucleotide complementary  
 CC to the complementary strand of a polynucleotide which comprises one of  
 CC the 5602 nucleotide sequences defined in the specification, where the  
 CC oligonucleotide comprises at least 15 nucleotides; or (b) a combination  
 CC of an oligonucleotide comprising a sequence complementary to the  
 CC complementary strand of a polynucleotide which comprises a 5'-end  
 CC sequence and an oligonucleotide comprising a sequence complementary to a  
 CC polynucleotide which comprises a 3'-end sequence, where the  
 CC oligonucleotide comprises at least 15 nucleotides and the combination of  
 CC the 5'-end sequence/3'-end sequence is selected from those defined in  
 CC the specification. The primer sets can be used in antisense therapy and  
 CC in gene therapy. The primers are useful for synthesizing polynucleotides,  
 CC particularly full-length cDNAs. The primers are also useful for the  
 CC detection and/or diagnosis of the abnormality of the proteins encoded by  
 CC the full-length cDNAs. The primers allow obtaining of the full-length

CC cDNAs easily without any specialised methods. AAH03166 to AAH13628 and  
 CC AAH13633 to AAH18742 represent human CDNA sequences; AAB92446 to  
 CC AAB95893 represent human amino acid sequences; and AAH13629 to AAH13632  
 CC represent oligonucleotides, all of which are used in the exemplification  
 of the present invention.

XX Sequence 353 AA:

Query Match 99.8%; Score 1200; DB 22; Length 353;  
 Best Local Similarity 99.6%; Pred. No. 9,8e-124;  
 Matches 227; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 MRYLLPSVVLGTAFTVYVLAAGVWRLLSAFLPARFYQALDDRLCYQSMVLEFFENYTG 60  
 DB 1 mryllpsvvlgtapcyvlawgwrllsalfparfyqalddrlcyvqsmvllffenytg 60  
 QY 61 VQILLYGDLPKKNENIITYLANHSTVDWIVADILAIRONALGHVRYVLEKGLKWLPLYGC 120  
 DB 61 vqilllygdldpkknentiitylanhgstvdwivadilairgnalghvryvlekgllwplygc 120  
 QY 121 YFAOHGITYVKRSKAKFNEKEMRNKLOSVDAGTPMTLVIFPEGTRINPEQTKVLSASQAF 180  
 DB 121 yfaohgityvkrsakfnekeemrnlqsyvdagtpmtylvifpegtrinypeqtkvlsasqaf 180  
 QY 181 AAOBGILAVLKHVLPRIKATHVAFDCKMKNYLDIYDVTVVYEGKDDGG 228  
 DB 181 aabgriavlkhlvlprikathvafdcmknlyldaiydvttvvyegkddgg 228

RESULT 10

AAH41377  
 ID AAM41377 standard; Protein; 372 AA.

XX AAM41377;

DT 22-OCT-2001 (first entry)

XX Human polypeptide SEQ ID NO 6308.

XX Human; nootropic; immunosuppressant; cytostatic; gene therapy; cancer;  
 KW peripheral nervous system; neuropathy; central nervous system; CNS;  
 KW Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;  
 KW amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;  
 KW chemokinetic; thrombolytic; drug screening; arthritis; inflammation;  
 KW leukaemia.

XX Homo sapiens.

XX WO200153312-A1.

XX 26-JUL-2001.

PF 26-DEC-2000; 2000WO-US34263.

XX 21-JAN-2000; 2000US-0488725.

PR 25-APR-2000; 2000US-0552317.

PR 09-JUL-2000; 2000US-0598042.

PR 19-JUL-2000; 2000US-0620312.

PR 03-AUG-2000; 2000US-0653450.

PR 14-SEP-2000; 2000US-0662191.

PR 19-OCT-2000; 2000US-0693036.

PR 29-NOV-2000; 2000US-0727344.

XX (HYSE-) HYSEQ INC.

XX Tang YF, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;  
 PI Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J;  
 PI Zhao QA, Zhou P, Goodrich R, Dzmanac RT;  
 XX WPI; 2001-442253/47.  
 DR N-PSDB; AAI60533.  
 XX

PT Novel nucleic acids and polypeptides, useful for treating disorders  
 PT such as central nervous system injuries -  
 XX Example 2; SEQ ID NO 6308; 10078bp; English.

XX The invention relates to human nucleic acids (AA157798-AA161369) and  
 CC the encoded polypeptides (AAM38642-AAH42213) with nootropic,  
 CC immunosuppressant and cytostatic activity. The polynucleotides are useful  
 CC in gene therapy. A composition containing a polypeptide or polynucleotide  
 CC of the invention may be used to treat diseases of the peripheral nervous  
 CC system, such as peripheral nervous injuries, peripheral neuropathy and  
 CC localized neuropathies and central nervous system diseases, such as  
 CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic  
 CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the  
 CC utilisation of the activities such as: Immune system suppression,  
 CC Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic  
 CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,  
 CC assays for receptor activity, arthritis and inflammation, leukaemias and  
 CC C.N.S disorders.  
 CC Note: The sequence data for this patent did not form part of the printed  
 specification.

XX Sequence 372 AA:

Query Match 99.1%; Score 1192; DB 22; Length 372;  
 Best Local Similarity 99.6%; Pred. No. 8.1e-123;  
 Matches 227; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MRYLLPSVVLGTAFTVYVLAAGVWRLLSAFLPARFYQALDDRLCYQSMVLEFFENYTG 60  
 DB 20 mryllpsvvlgtapcyvlawgwrllsalfparfyqalddrlcyvqsmvllffenytg 79  
 QY 61 VQILLYGDLPKKNENIITYLANHSTVDWIVADILAIRONALGHVRYVLEKGLKWLPLYGC 120  
 DB 80 vqilllygdldpkknentiitylanhgstvdwivadilairgnalghvryvlekgllwplygw 139  
 QY 121 YFAOHGITYVKRSKAKFNEKEMRNKLOSVDAGTPMTLVIFPEGTRINPEQTKVLSASQAF 180  
 DB 140 yfaohgityvkrsakfnekeemrnlqsyvdagtpmtylvifpegtrinypeqtkvlsasqaf 199  
 QY 181 AAOBGILAVLKHVLPRIKATHVAFDCKMKNYLDIYDVTVVYEGKDDGG 228  
 DB 200 aabgriavlkhlvlprikathvafdcmknlyldaiydvttvvyegkddgg 247

RESULT 11

AAH36741  
 ID AAY36741 standard; Protein; 354 AA.

XX AAY36741;

DT 27-SEP-1999 (first entry)

XX Mouse PGI protein sequence.

DE

KW PGI gene; diallelic marker; mouse; PSA; PGI-related diallelic marker;  
 KW cancer; prostate cancer; diagnosis; therapy; prostate specific antigen.

XX Mus musculus.

PN WO9932644-A2.

XX 01-JUL-1999.

PF 22-DEC-1998; 98WO-IB02133.

PR 09-SEP-1998; 98US-0099658.

PR 22-DEC-1997; 97US-0996306.

XX (GEST ) GENSET.  
 PA Blumenfeld M, Bougueleret L, Chumakov I, Cohen D;  
 PI

XX WPI: 1999-405178/34.  
DR N-PSDB: AAZ00928, AAZ01027.  
XX  
PT Use of a prostate cancer associated gene and biallelic markers  
PT derived from it  
XX  
PS Claim 7; Page 215; 385pp; English.  
XX  
CS The invention relates to a mammalian pgl gene and protein, and a set of  
CC pgl biallelic markers. The pgl polynucleotide and biallelic markers are  
CC used in a hybridisation assay, a sequencing assay, or in an  
CC allele-specific amplification assay for determining the identity of a  
CC nucleotide at a pgl-related biallelic marker. The methods can be used to  
CC detect and to assess the risk of developing cancer or prostate cancer.  
CC Early-stage diagnosis of prostate cancer relies on prostate specific  
CC antigen (PSA) dosage. However, the effectiveness of this is limited due  
CC to its inability to discriminate between malignant and non-malignant  
CC affections of the organ. A need exists for both a reliable diagnostic  
CC procedure which would enable early-stage diagnosis, and for preventative  
CC and curative treatments of the disease. The pgl gene can be used for  
CC detection of prostate cancer, and the risk of developing it in the  
CC future, and can also be used to determine therapies for the disease.  
XX  
SQ Sequence 354 AA;

Query Match 86.5%; Score 1040; DB 20; Length 354;  
Best Local Similarity 85.9%; Pred. No. 4.4e-106;  
Matches 195; Conservative 19; Mismatches 13; Indels 0; Gaps 0;  
QY 1 MRYLLPSVLLGTAPTYVLANGVWRLLSAFLPARFYQALDRLCYVQSMVLPFFENYTG 60  
DB 1 mryllpsvlllgsaptyllawclwrvisalmparfyqvdrrlcyvqsmvllffenytg 60  
QY 61 VQILLYGDLPKNKENIYILANHOSTVDWIVADILAIRONALGHWRYVLEKGLKMLPLYGC 120  
DB 61 vqilllygdlpknkennyilanhgstvdwivadilaarqdalghvrlyvllkdklkwlplygf 120  
QY 121 YFAOHGCIYVRSKAFKFNKEMRNKLSQSYVDAGTPMYLVIFPEGTRVNEQTKVLSASQAF 180  
DB 121 yfaohgciyvrskafkfnkemrnsklsqsyvndagtpmylvifpegtrvneatkllsasqaf 180  
QY 181 AAOGLAVLKHVLTFRKATVAVFDCMKKNYLDALYDVTVYEGKDDG 227  
DB 181 aagrglavlkhnvltfrkathvafdcmskshldalydvrvlyegnekgy 227

RESULT 12  
AAV36753  
ID AAV36753 standard; Protein: 185 AA.  
XX  
AC AAV36753;  
XX  
DT 27-SEP-1999 (first entry)  
XX  
DE pgl splice variant allele protein sequence.  
XX  
KW pgl gene; biallelic marker; human; PSA; pgl-related biallelic marker;  
KM cancer; prostate cancer; diagnosis; therapy; prostate specific antigen.  
XX  
OS Homo sapiens.  
XX  
PN WO9932644-A2.  
XX  
PD 01-JUL-1999.  
XX  
PF 22-DEC-1998; 98WO-IB02133.  
XX  
PR 09-SEP-1998; 98US-0099658.  
XX  
PR 22-DEC-1997; 97US-0996306.  
XX  
PA (GEST ) GENSET.

XX Blumenfeld M, Bougueleret L, Chumakov I, Cohen D;  
PI WPI: 1999-405178/34.  
XX  
DR N-PSDB: AAZ00979.  
XX  
PT Use of a prostate cancer associated gene and biallelic markers  
PT derived from it  
XX  
PS Claim 7; Page 255; 385pp; English.  
XX  
CS The invention relates to a mammalian pgl gene and protein, and a set of  
CC pgl biallelic markers. The pgl polynucleotide and biallelic markers are  
CC used in a hybridisation assay, a sequencing assay, or in an  
CC allele-specific amplification assay for determining the identity of a  
CC nucleotide at a pgl-related biallelic marker. The methods can be used to  
CC detect and to assess the risk of developing cancer or prostate cancer.  
CC Early-stage diagnosis of prostate cancer relies on prostate specific  
CC antigen (PSA) dosage. However, the effectiveness of this is limited due  
CC to its inability to discriminate between malignant and non-malignant  
CC affections of the organ. A need exists for both a reliable diagnostic  
CC procedure which would enable early-stage diagnosis, and for preventative  
CC and curative treatments of the disease. The pgl gene can be used for  
CC detection of prostate cancer, and the risk of developing it in the  
CC future, and can also be used to determine therapies for the disease.  
XX  
SQ Sequence 185 AA;

Query Match 81.1%; Score 976; DB 20; Length 185;  
Best Local Similarity 100.0%; Pred. No. 2.1e-99;  
Matches 185; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 MRYLLPSVLLGTAPTYVLANGVWRLLSAFLPARFYQALDRLCYVQSMVLPFFENYTG 60  
DB 1 mryllpsvlllgtaptyllawgvwrllsafiparfyaldrlycvqsmvllffenytg 60  
QY 61 VQILLYGDLPKNKENIYILANHOSTVDWIVADILAIRONALGHWRYVLEKGLKMLPLYGC 120  
DB 61 vqilllygdlpknkennyilanhgstvdwivadilaarqnalghvrlyvllkdklkwlplygc 120  
QY 121 YFAOHGCIYVRSKAFKFNKEMRNKLSQSYVDAGTPMYLVIFPEGTRVNEQTKVLSASQAF 180  
DB 121 yfaohgciyvrskafkfnkemrnsklsqsyvndagtpmylvifpegtrvneatkllsasqaf 180  
QY 181 AAOGRG 185  
DB 181 aagrg 185

RESULT 13  
AAV36751  
ID AAV36751 standard; Protein: 315 AA.  
XX  
AC AAV36751;  
XX  
DT 27-SEP-1999 (first entry)  
XX  
DE pgl splice variant allele protein sequence.  
XX  
KW pgl gene; biallelic marker; human; PSA; pgl-related biallelic marker;  
KM cancer; prostate cancer; diagnosis; therapy; prostate specific antigen.  
XX  
OS Homo sapiens.  
XX  
PN WO9932644-A2.  
XX  
PD 01-JUL-1999.  
XX  
PF 22-DEC-1998; 98WO-IB02133.  
XX  
PR 09-SEP-1998; 98US-0099658.  
XX  
PR 22-DEC-1997; 97US-0996306.

XX (GEST ) GENSET.  
PA Blumenfeld M, Bougueleret L, Chumakov I, Cohen D;  
XX WPI: 1999-405178/34.  
DR N-PSDB: AA200977.  
XX  
PT Use of a prostate cancer associated gene and biallelic markers  
XX derived from it  
XX  
PS Claim 7; Page 253-254; 385pp; English.  
XX  
CC The invention relates to a mammalian PGI gene and protein, and a set of  
CC PGI biallelic markers. The PGI polynucleotide and biallelic markers are  
CC used in a hybridisation assay, a sequencing assay, or in an  
CC allele-specific amplification assay for determining the identity of a  
CC nucleotide at a PGI-related biallelic marker. The methods can be used to  
CC detect and to assess the risk of developing cancer or prostate cancer.  
CC Early-stage diagnosis of prostate cancer relies on prostate specific  
CC antigen (PSA) dosage. However, the effectiveness of this is limited due  
CC to its inability to discriminate between malignant and non-malignant  
CC affections of the organ. A need exists for both a reliable diagnostic  
CC procedure which would enable early-stage diagnosis, and for preventative  
CC and curative treatments of the disease. The PGI gene can be used for  
CC detection of prostate cancer, and the risk of developing it in the  
CC future, and can also be used to determine therapies for the disease.  
XX  
SQ Sequence 315 AA;

Query Match 81.1%; Score 976; DB 20; Length 315;  
Best Local Similarity 100.0%; Pred. No. 4.3e-99;  
Matches 185; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 MRYLLPSVVLGTAPTYVLAMGWRLISAFLPARFQALDDRLCYQSVLFFEFNYTG 60  
DB 1 mryllpsvvlgtaptyvlamgwrlisafllparfygalddrlycyqsmvlffefnytg 60  
QY 61 VQILLXGDLPRKKNENTIIYLANHOSTVDWIVADILAIROMALGHVRYVLEKGLKWLPLYGC 120  
DB 61 vqillxgdlprkknentiiylanhostvdwivadilairgnalghvryvlekgllkwlplygc 120  
QY 121 YFAOHGCIYKRSKAKFNKEMRKLOSVDAGTPMYLVIFPEGTRNPEOTKVLASQAF 180  
DB 121 yfaohgciykrskakfnkemrklqsyvdagtpmylvifpegtrypetkvlisasqaf 180  
QY 181 AAORG 185  
DB 181 aagrg 185

RESULT 14  
AAY36752  
ID AAY36752 standard; Protein: 300 AA.  
XX  
AC AAY36752;  
XX  
DT 27-SEP-1999 (first entry)  
XX  
DE PGI splice variant allele protein sequence.  
XX  
KW PGI gene; biallelic marker; human; PSA; PGI-related biallelic marker;  
KW cancer; prostate cancer; diagnosis; therapy; prostate specific antigen.  
XX Homo sapiens.  
OS  
PN WO932644-A2.  
XX  
PD 01-JUL-1999.  
XX  
PF 22-DEC-1998; 98WO-IB02133.  
XX

PR 09-SEP-1998; 98US-0099658.  
PR 22-DEC-1997; 97US-0996306.  
XX  
PA (GEST ) GENSET.  
XX  
PI Blumenfeld M, Bougueleret L, Chumakov I, Cohen D;  
XX WPI: 1999-405178/34.  
DR N-PSDB: AA200978.  
XX  
PT Use of a prostate cancer associated gene and biallelic markers  
XX derived from it  
XX  
PS Claim 7; Page 254-255; 385pp; English.  
XX  
CC The invention relates to a mammalian PGI gene and protein, and a set of  
CC PGI biallelic markers. The PGI polynucleotide and biallelic markers are  
CC used in a hybridisation assay, a sequencing assay, or in an  
CC allele-specific amplification assay for determining the identity of a  
CC nucleotide at a PGI-related biallelic marker. The methods can be used to  
CC detect and to assess the risk of developing cancer or prostate cancer.  
CC Early-stage diagnosis of prostate cancer relies on prostate specific  
CC antigen (PSA) dosage. However, the effectiveness of this is limited due  
CC to its inability to discriminate between malignant and non-malignant  
CC affections of the organ. A need exists for both a reliable diagnostic  
CC procedure which would enable early-stage diagnosis, and for preventative  
CC and curative treatments of the disease. The PGI gene can be used for  
CC detection of prostate cancer, and the risk of developing it in the  
CC future, and can also be used to determine therapies for the disease.  
XX  
SQ Sequence 300 AA;

Query Match 80.6%; Score 970; DB 20; Length 300;  
Best Local Similarity 100.0%; Pred. No. 1.8e-98;  
Matches 184; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 MRYLLPSVVLGTAPTYVLAMGWRLISAFLPARFQALDDRLCYQSVLFFEFNYTG 60  
DB 1 mryllpsvvlgtaptyvlamgwrlisafllparfygalddrlycyqsmvlffefnytg 60  
QY 61 VQILLXGDLPRKKNENTIIYLANHOSTVDWIVADILAIROMALGHVRYVLEKGLKWLPLYGC 120  
DB 61 vqillxgdlprkknentiiylanhostvdwivadilairgnalghvryvlekgllkwlplygc 120  
QY 121 YFAOHGCIYKRSKAKFNKEMRKLOSVDAGTPMYLVIFPEGTRNPEOTKVLASQAF 180  
DB 121 yfaohgciykrskakfnkemrklqsyvdagtpmylvifpegtrypetkvlisasqaf 180  
QY 181 AAOR 184  
DB 181 aagr 184

RESULT 15  
AAY36744  
ID AAY36744 standard; Protein: 291 AA.  
XX  
AC AAY36744;  
XX  
DT 27-SEP-1999 (first entry)  
XX  
DE PGI splice variant allele protein sequence.  
XX  
KW PGI gene; biallelic marker; human; PSA; PGI-related biallelic marker;  
KW cancer; prostate cancer; diagnosis; therapy; prostate specific antigen.  
XX Homo sapiens.  
OS  
PN WO932644-A2.  
XX  
PD 01-JUL-1999.  
XX

PF 22-DEC-1998; 98MO-IB02133.  
 XX  
 PR 09-SEP-1998; 98US-0099658.  
 PR 22-DEC-1997; 97US-0096306.  
 XX  
 PA (GEST ) GENSET.  
 XX  
 PI Blumenfeld M, Bougueleret L, Chumakov I, Cohen D;  
 XX  
 DR WPI; 1999-405178/34.  
 DR N-PSDB; AA200970.  
 XX  
 PT Use of a prostate cancer associated gene and biallelic markers  
 derived from it  
 XX  
 PS Claim 7; Page 250-251; 385pp; English.  
 XX  
 CC The invention relates to a mammalian Pcl gene and protein, and a set of  
 CC Pcl biallelic markers. The Pcl polynucleotide and biallelic markers are  
 CC used in a hybridisation assay, a sequencing assay, or in an  
 CC allele-specific amplification assay for determining the identity of a  
 CC nucleotide at a Pcl-related biallelic marker. The methods can be used to  
 CC detect and to assess the risk of developing cancer or prostate cancer.  
 CC Early-stage diagnosis of prostate cancer relies on prostate specific  
 CC antigen (PSA) dosage. However, the effectiveness of this is limited due  
 CC to its inability to discriminate between malignant and non-malignant  
 CC affections of the organ. A need exists for both a reliable diagnostic  
 CC procedure which would enable early-stage diagnosis, and for preventative  
 CC and curative treatments of the disease. The Pcl gene can be used for  
 CC detection of prostate cancer, and the risk of developing it in the  
 CC future, and can also be used to determine therapies for the disease.  
 CC  
 XX  
 SQ Sequence 291 AA;

Query Match 69.0%; Score 830; DB 20; Length 291;  
 Best Local Similarity 72.8%; Pred. No. 4.9e-83;  
 Matches 166; Conservative 0; Mismatches 0; Indels 62; Gaps 1;  
 QY 1 MRYLLPSVVLGTAPTYVVLAWGVRLLSAFLPARFYQALDDRLCYVQSMVLPFFENYTG 60  
 Db 1 MRYLLPSVVLGTAPTYVVLAWGVRLLSAFLPARFYQALDDRLCYVQSMVLPFFENYTG 60  
 QY 61 VOILLYGDLPKKNENITYLANHGSTVDWIVADILAIRONALGHVRYVLEKGLKWLPLXGC 120  
 Db 61 v----- 61  
 QY 121 YFAHQGCIYVKRSKFNKEMRNKLOSVDAGTPMYLVIFPEGTRYNPEQTKVLSASQAF 180  
 Db 62 ---qhg9lyvkrasakfnekemrnlqsyvdagtpmylvlfpegtrynpegtkvlsasqaf 118  
 QY 181 AAGRGLAVLKHVLTPIKATVAFDCMKKNYLDALYDVTVVYEGKDDG 228  
 Db 119 aagrglavlkhvltprikatvafdcmkknlyldalydvttvyegkddg 166

Search completed: August 28, 2002, 11:07:35  
 Job time: 479 sec



GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 28, 2002, 11:14:11 ; Search time 306.96 Seconds

(without alignments)  
261.440 Million cell updates/sec

Title: US-09-853-526-70

Perfect score: 1203  
Sequence: 1 MRYLPSVVLGTAPIYVLA.....NYLDAIVYVYEGKDDG 228

Scoring table:

BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 3502263 segs, 351980561 residues

Total number of hits satisfying chosen parameters: 3502263

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Pending Patents\_AA\_Main:\*  
1: /cgn2\_6/ptodata/2/paa/US085\_COMB.pep:\*  
2: /cgn2\_6/ptodata/2/paa/US06\_COMB.pep:\*  
3: /cgn2\_6/ptodata/2/paa/US07\_COMB.pep:\*  
4: /cgn2\_6/ptodata/2/paa/US08\_COMB.pep:\*  
5: /cgn2\_6/ptodata/2/paa/US081\_COMB.pep:\*  
6: /cgn2\_6/ptodata/2/paa/US082\_COMB.pep:\*  
7: /cgn2\_6/ptodata/2/paa/US083\_COMB.pep:\*  
8: /cgn2\_6/ptodata/2/paa/US084\_COMB.pep:\*  
9: /cgn2\_6/ptodata/2/paa/US085\_COMB.pep:\*  
10: /cgn2\_6/ptodata/2/paa/US086\_COMB.pep:\*  
11: /cgn2\_6/ptodata/2/paa/US087\_COMB.pep:\*  
12: /cgn2\_6/ptodata/2/paa/US088\_COMB.pep:\*  
13: /cgn2\_6/ptodata/2/paa/US089\_COMB.pep:\*  
14: /cgn2\_6/ptodata/2/paa/US090\_COMB.pep:\*  
15: /cgn2\_6/ptodata/2/paa/US091\_COMB.pep:\*  
16: /cgn2\_6/ptodata/2/paa/US092\_COMB.pep:\*  
17: /cgn2\_6/ptodata/2/paa/US093\_COMB.pep:\*  
18: /cgn2\_6/ptodata/2/paa/US094\_COMB.pep:\*  
19: /cgn2\_6/ptodata/2/paa/US095\_COMB.pep:\*  
20: /cgn2\_6/ptodata/2/paa/US096\_COMB.pep:\*  
21: /cgn2\_6/ptodata/2/paa/US097\_COMB.pep:\*  
22: /cgn2\_6/ptodata/2/paa/US098\_COMB.pep:\*  
23: /cgn2\_6/ptodata/2/paa/US099\_COMB.pep:\*  
24: /cgn2\_6/ptodata/2/paa/US100\_COMB.pep:\*  
25: /cgn2\_6/ptodata/2/paa/US101\_COMB.pep:\*  
26: /cgn2\_6/ptodata/2/paa/US60\_COMB.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1203	100.0	228	US-09-853-526-70	Sequence 70, Appl
2	1203	100.0	228	US-09-901-484A-70	Sequence 70, Appl
3	1203	100.0	228	US-60-099-658-70	Sequence 70, Appl
4	1203	100.0	353	PCT-US01-01431-59	Sequence 59, Appl
5	1203	100.0	353	PCT-US01-01431-79	Sequence 79, Appl
6	1203	100.0	353	PCT-US01-11988-1474	Sequence 1474, Ap
7	1203	100.0	353	PCT-US01-11988-1475	Sequence 1475, Ap

8	1203	100.0	353	18	US-09-488-725A-2736	Sequence 2736, Ap
9	1203	100.0	353	22	US-09-833-245-1474	Sequence 1474, Ap
10	1203	100.0	353	22	US-09-833-245-1475	Sequence 1475, Ap
11	1203	100.0	353	23	US-09-853-526-4	Sequence 4, Appl1
12	1203	100.0	353	23	US-09-901-484A-4	Sequence 59, Appl
13	1203	100.0	353	23	US-09-915-582-59	Sequence 79, Appl
14	1203	100.0	353	23	US-09-915-582-79	Sequence 4, Appl1
15	1203	100.0	353	26	US-60-099-658-4	Sequence 1, Appl1
16	1203	100.0	364	18	US-09-436-919-1	Sequence 7, Appl1
17	1203	100.0	364	22	US-09-817-910-7	Sequence 5, Appl1
18	1203	100.0	364	22	US-09-853-526-5	Sequence 5, Appl1
19	1203	100.0	364	23	US-09-901-484A-5	Sequence 5, Appl1
20	1203	100.0	364	26	US-60-099-658-5	Sequence 6308, Ap
21	1192	99.1	372	18	US-09-488-725A-6308	Sequence 110, App
22	1066.5	88.7	450	26	US-60-245-222-110	Sequence 309, App
23	1043.5	86.7	1032	26	US-60-212-413-309	Sequence 248, App
24	1043.5	86.7	1032	26	US-60-229-518-248	Sequence 74, Appl
25	1040	86.5	354	22	US-09-853-526-74	Sequence 74, Appl
26	1040	86.5	354	23	US-09-901-484A-74	Sequence 74, Appl
27	1040	86.5	354	26	US-60-099-658-74	Sequence 136, App
28	976	81.1	185	22	US-09-853-526-136	Sequence 136, App
29	976	81.1	185	23	US-09-901-484A-136	Sequence 134, App
30	976	81.1	315	22	US-09-901-484A-134	Sequence 135, App
31	976	81.1	315	23	US-09-853-526-135	Sequence 135, App
32	970	80.6	300	22	US-09-901-484A-135	Sequence 127, App
33	970	80.6	300	23	US-09-853-526-127	Sequence 127, App
34	830	69.0	291	22	US-09-901-484A-127	Sequence 133, App
35	830	69.0	291	23	US-09-853-526-133	Sequence 133, App
36	820	68.2	182	22	US-09-901-484A-133	Sequence 47, Appl
37	820	68.2	182	23	PCT-US01-01327-47	Sequence 128, App
38	759	63.1	269	1	US-09-853-526-128	Sequence 128, App
39	657	54.6	261	23	US-09-901-484A-128	Sequence 126, App
40	657	54.6	261	23	US-09-853-526-126	Sequence 126, App
41	597	49.6	238	22	US-09-901-484A-126	Sequence 5, Appl1
42	597	49.6	238	23	US-09-436-919-5	Sequence 132, App
43	594	48.4	176	18	US-09-853-526-132	Sequence 132, App
44	449	37.3	97	22	US-09-901-484A-132	Sequence 132, App
45	449	37.3	97	23	US-09-901-484A-132	Sequence 132, App

#### ALIGNMENTS

RESULT 1  
US-09-853-526-70  
; Sequence 70, Application US/09853526  
; GENERAL INFORMATION:  
; APPLICANT: Cohen, Daniel  
; APPLICANT: Blumenfeld, Marta  
; APPLICANT: Ilyia, Chumakov  
; APPLICANT: Bouguetel, Lydie  
; TITLE OF INVENTION: PROSTATE CANCER GENE  
; FILE REFERENCE: GENSET.18CP1CP  
; CURRENT APPLICATION NUMBER: US/09/853,526  
; CURRENT FILING DATE: 2001-05-11  
; PRIOR APPLICATION NUMBER: 09/338, 907  
; PRIOR FILING DATE: 1999-06-23  
; PRIOR APPLICATION NUMBER: 08/996, 306  
; PRIOR FILING DATE: 1997-12-22  
; PRIOR APPLICATION NUMBER: 60/099, 658  
; PRIOR FILING DATE: 1998-09-09  
; PRIOR APPLICATION NUMBER: 09/218, 207  
; PRIOR FILING DATE: 1998-12-22  
; NUMBER OF SEQ ID NOS: 578  
; SOFTWARE: Patent.pm  
; SEQ ID NO 70  
; LENGTH: 228  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-853-526-70  
Query Match 100.0%; Score 1203; DB 22; Length 228;

Best Local Similarity 100.0%; Pred. No. 7.7e-120;  
Matches 228; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY 1 MRYLLPSVLLGTAPTYVLAAGWVRLSAFLPARFYQALDDRLCYVOSMVLFFFEENTYG 60
    |||
DB 1 MRYLLPSVLLGTAPTYVLAAGWVRLSAFLPARFYQALDDRLCYVOSMVLFFFEENTYG 60
    |||
QY 61 VOILLYGDLPKKNENITYLANHOSYVDWIVADILAIKRONALGHVRYVKEGLKMLPLXGC 120
    |||
DB 61 VOILLYGDLPKKNENITYLANHOSYVDWIVADILAIKRONALGHVRYVKEGLKMLPLXGC 120
    |||
QY 121 YFAOHGIIYVKSAAFKNEKEMRNKLOSIVDAGTPMYLVIFPEGTRYNPEQTKVLSASQAF 180
    |||
DB 121 YFAOHGIIYVKSAAFKNEKEMRNKLOSIVDAGTPMYLVIFPEGTRYNPEQTKVLSASQAF 180
    |||
QY 181 AAORGLAVLKHVLTPIKATHVAFDCMKNYLDIYDVTVYVEGKDDGG 228
    |||
DB 181 AAORGLAVLKHVLTPIKATHVAFDCMKNYLDIYDVTVYVEGKDDGG 228
    |||
```

## RESULT 2

US-09-901-484A-70

; Sequence 70, Application US/09901484A  
; GENERAL INFORMATION:

; APPLICANT: Cohen, Daniel  
; APPLICANT: Blumenfeld, Marta  
; APPLICANT: Chumakov, Ilya  
; APPLICANT: Bougueleret, Lydie  
; TITLE OF INVENTION: Prostate Cancer Gene  
; FILE REFERENCE: GEN-T11XC3D2  
; CURRENT APPLICATION NUMBER: US/09/901,484A  
; CURRENT FILING DATE: 2001-07-09  
; PRIOR APPLICATION NUMBER: US 08/996,306  
; PRIOR FILING DATE: 1997-12-22  
; PRIOR APPLICATION NUMBER: US 60/099,658  
; PRIOR FILING DATE: 1998-09-09  
; PRIOR APPLICATION NUMBER: US 09/218,207  
; PRIOR FILING DATE: 1998-12-22  
; PRIOR APPLICATION NUMBER: US 09/338,907  
; PRIOR FILING DATE: 1999-06-23  
; PRIOR APPLICATION NUMBER: US 09/853,526  
; PRIOR FILING DATE: 2001-05-11  
; NUMBER OF SEQ ID NOS: 578  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 70  
; LENGTH: 228  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: misc.feature  
; LOCATION: (5208)..(5208)  
; OTHER INFORMATION: n = a, c, g, or t.

US-09-901-484A-70

Query Match 100.0%; Score 1203; DB 23; Length 228;  
Best Local Similarity 100.0%; Pred. No. 7.7e-120;  
Matches 228; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY 1 MRYLLPSVLLGTAPTYVLAAGWVRLSAFLPARFYQALDDRLCYVOSMVLFFFEENTYG 60
    |||
DB 1 MRYLLPSVLLGTAPTYVLAAGWVRLSAFLPARFYQALDDRLCYVOSMVLFFFEENTYG 60
    |||
QY 61 VOILLYGDLPKKNENITYLANHOSYVDWIVADILAIKRONALGHVRYVKEGLKMLPLXGC 120
    |||
DB 61 VOILLYGDLPKKNENITYLANHOSYVDWIVADILAIKRONALGHVRYVKEGLKMLPLXGC 120
    |||
QY 121 YFAOHGIIYVKSAAFKNEKEMRNKLOSIVDAGTPMYLVIFPEGTRYNPEQTKVLSASQAF 180
    |||
DB 121 YFAOHGIIYVKSAAFKNEKEMRNKLOSIVDAGTPMYLVIFPEGTRYNPEQTKVLSASQAF 180
    |||
QY 181 AAORGLAVLKHVLTPIKATHVAFDCMKNYLDIYDVTVYVEGKDDGG 228
    |||
```

DB 181 AAORGLAVLKHVLTPIKATHVAFDCMKNYLDIYDVTVYVEGKDDGG 228

## RESULT 3

US-60-099-658-70

; Sequence 70, Application US/60099658  
; GENERAL INFORMATION:

; APPLICANT: Cohen, Daniel  
; APPLICANT: Chumakov, Ilya  
; APPLICANT: Blumenfeld, Marta  
; APPLICANT: Bougueleret, Lydie  
; TITLE OF INVENTION: Prostate cancer gene  
; NUMBER OF SEQUENCES: 99  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Knobbe, Martens, Olson & Bear  
; STREET: 501 West Broadway  
; CITY: San Diego  
; STATE: California  
; COUNTRY: USA  
; ZIP: 92101-3505

; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: Win95  
; SOFTWARE: Word

; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/60/099,658  
; FILING DATE:  
; CLASSIFICATION:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Israelson, Ned A.  
; REGISTRATION NUMBER: 29,655  
; REFERENCE/DOCKET NUMBER: GENSET.018APR  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (619) 235-8550  
; TELEFAX: (619) 235-0176

; INFORMATION FOR SEQ ID NO: 70:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 228 amino acids  
; TYPE: AMINO ACID  
; STRANDEDNESS: SINGLE  
; TOPOLOGY: LINEAR  
; MOLECULE TYPE: Protein  
; ORIGINAL SOURCE:  
; ORGANISM: Homo sapiens

; US-60-099-658-70

Query Match 100.0%; Score 1203; DB 26; Length 228;  
Best Local Similarity 100.0%; Pred. No. 7.7e-120;  
Matches 228; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY 1 MRYLLPSVLLGTAPTYVLAAGWVRLSAFLPARFYQALDDRLCYVOSMVLFFFEENTYG 60
    |||
DB 1 MRYLLPSVLLGTAPTYVLAAGWVRLSAFLPARFYQALDDRLCYVOSMVLFFFEENTYG 60
    |||
QY 61 VOILLYGDLPKKNENITYLANHOSYVDWIVADILAIKRONALGHVRYVKEGLKMLPLXGC 120
    |||
DB 61 VOILLYGDLPKKNENITYLANHOSYVDWIVADILAIKRONALGHVRYVKEGLKMLPLXGC 120
    |||
QY 121 YFAOHGIIYVKSAAFKNEKEMRNKLOSIVDAGTPMYLVIFPEGTRYNPEQTKVLSASQAF 180
    |||
DB 121 YFAOHGIIYVKSAAFKNEKEMRNKLOSIVDAGTPMYLVIFPEGTRYNPEQTKVLSASQAF 180
    |||
QY 181 AAORGLAVLKHVLTPIKATHVAFDCMKNYLDIYDVTVYVEGKDDGG 228
    |||
DB 181 AAORGLAVLKHVLTPIKATHVAFDCMKNYLDIYDVTVYVEGKDDGG 228
    |||
```

## RESULT 4

PCT-US01-01431-59

; Sequence 59, Application PC/TUS0101431  
; GENERAL INFORMATION:



NUMBER OF SEQ ID NOS: 2267  
SOFTWARE: Patent Ver. 2.1  
SEQ ID NO: 1475  
LENGTH: 353  
TYPE: PRT  
ORGANISM: Homo sapiens  
PCT-US01-11988-1475

Query Match 100.0%; Score 1203; DB 1; Length 353;  
Best Local Similarity 100.0%; Pred. No. 1.4e-119;  
Matches 228; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MRYLLPSVLLGTAFTYVLANGVWRLSAPLPAFYQALDRRLCYQSMVLPFFENYTG 60  
DB 1 MRYLLPSVLLGTAFTYVLANGVWRLSAPLPAFYQALDRRLCYQSMVLPFFENYTG 60  
QY 61 VOILLYGDLPKKNENIYLANHSTVDMIVADILAIRONALGHVRYVLEKGLKWLPLYGC 120  
DB 61 VOILLYGDLPKKNENIYLANHSTVDMIVADILAIRONALGHVRYVLEKGLKWLPLYGC 120  
QY 121 YFAOHGCIYVRSKAFNEKEKRNKLOSVDAGTMYLVIPEEGTRYNPEQTKVLSASQAF 180  
DB 121 YFAOHGCIYVRSKAFNEKEKRNKLOSVDAGTMYLVIPEEGTRYNPEQTKVLSASQAF 180  
QY 181 AAQGLAVLKHVLTPIKATHVAFDCKMKNYLDATYDVTVYEGKDDG 228  
DB 181 AAQGLAVLKHVLTPIKATHVAFDCKMKNYLDATYDVTVYEGKDDG 228

RESULT 8  
US-09-488-725A-2736  
Sequence 2736, Application US/09488725A  
GENERAL INFORMATION:  
APPLICANT: Hyseq Inc  
TITLE OF INVENTION: Novel Nucleic Acid and Polypeptides  
FILE REFERENCE: 784PLPCT  
CURRENT APPLICATION NUMBER: US/09/488, 725A  
CURRENT FILING DATE: 2000-12-22  
PRIOR APPLICATION NUMBER: US/09/488, 725  
PRIOR FILING DATE: 2000-01-21  
PRIOR APPLICATION NUMBER: US09/552, 317  
PRIOR FILING DATE: 2000-04-25  
PRIOR APPLICATION NUMBER: US09/598, 042  
PRIOR FILING DATE: 2000-06-20  
PRIOR APPLICATION NUMBER: US09/620, 312  
PRIOR FILING DATE: 2000-07-19  
PRIOR APPLICATION NUMBER: US09/653, 450  
PRIOR FILING DATE: 2000-08-31  
PRIOR APPLICATION NUMBER: US09/662, 191  
PRIOR FILING DATE: 2000-09-14  
PRIOR APPLICATION NUMBER: US09/693, 036  
PRIOR FILING DATE: 2000-10-19  
PRIOR APPLICATION NUMBER: US09/727, 344  
PRIOR FILING DATE: 2000-11-29  
NUMBER OF SEQ ID NOS: 7144  
SOFTWARE: PL\_FL\_genes\_b Versions 1.0  
SEQ ID NO 2736  
LENGTH: 354  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-09-488-725A-2736

Query Match 100.0%; Score 1203; DB 18; Length 353;  
Best Local Similarity 100.0%; Pred. No. 1.4e-119;  
Matches 228; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MRYLLPSVLLGTAFTYVLANGVWRLSAPLPAFYQALDRRLCYQSMVLPFFENYTG 60  
DB 1 MRYLLPSVLLGTAFTYVLANGVWRLSAPLPAFYQALDRRLCYQSMVLPFFENYTG 60  
QY 61 VOILLYGDLPKKNENIYLANHSTVDMIVADILAIRONALGHVRYVLEKGLKWLPLYGC 120  
DB 61 VOILLYGDLPKKNENIYLANHSTVDMIVADILAIRONALGHVRYVLEKGLKWLPLYGC 120

DB 61 VOILLYGDLPKKNENIYLANHSTVDMIVADILAIRONALGHVRYVLEKGLKWLPLYGC 120  
QY 121 YFAOHGCIYVRSKAFNEKEKRNKLOSVDAGTMYLVIPEEGTRYNPEQTKVLSASQAF 180  
DB 121 YFAOHGCIYVRSKAFNEKEKRNKLOSVDAGTMYLVIPEEGTRYNPEQTKVLSASQAF 180  
QY 181 AAQGLAVLKHVLTPIKATHVAFDCKMKNYLDATYDVTVYEGKDDG 228  
DB 181 AAQGLAVLKHVLTPIKATHVAFDCKMKNYLDATYDVTVYEGKDDG 228

RESULT 9  
US-09-833-245-1474  
Sequence 1474, Application US/09833245  
GENERAL INFORMATION:  
APPLICANT: Human Genome Sciences, Inc.  
TITLE OF INVENTION: Albumin Fusion Proteins  
FILE REFERENCE: PFS46PCT  
CURRENT APPLICATION NUMBER: US/09/833, 245  
CURRENT FILING DATE: 2001-04-12  
PRIOR APPLICATION NUMBER: 60/229, 358  
PRIOR FILING DATE: 2000-04-12  
PRIOR APPLICATION NUMBER: 60/256, 931  
PRIOR FILING DATE: 2000-12-21  
PRIOR APPLICATION NUMBER: 60/199, 384  
PRIOR FILING DATE: 2000-04-25  
NUMBER OF SEQ ID NOS: 2267  
SOFTWARE: Patent Ver. 2.1  
SEQ ID NO 1474  
LENGTH: 353  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-09-833-245-1474

Query Match 100.0%; Score 1203; DB 22; Length 353;  
Best Local Similarity 100.0%; Pred. No. 1.4e-119;  
Matches 228; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MRYLLPSVLLGTAFTYVLANGVWRLSAPLPAFYQALDRRLCYQSMVLPFFENYTG 60  
DB 1 MRYLLPSVLLGTAFTYVLANGVWRLSAPLPAFYQALDRRLCYQSMVLPFFENYTG 60  
QY 61 VOILLYGDLPKKNENIYLANHSTVDMIVADILAIRONALGHVRYVLEKGLKWLPLYGC 120  
DB 61 VOILLYGDLPKKNENIYLANHSTVDMIVADILAIRONALGHVRYVLEKGLKWLPLYGC 120  
QY 121 YFAOHGCIYVRSKAFNEKEKRNKLOSVDAGTMYLVIPEEGTRYNPEQTKVLSASQAF 180  
DB 121 YFAOHGCIYVRSKAFNEKEKRNKLOSVDAGTMYLVIPEEGTRYNPEQTKVLSASQAF 180  
QY 181 AAQGLAVLKHVLTPIKATHVAFDCKMKNYLDATYDVTVYEGKDDG 228  
DB 181 AAQGLAVLKHVLTPIKATHVAFDCKMKNYLDATYDVTVYEGKDDG 228

RESULT 10  
US-09-833-245-1475  
Sequence 1475, Application US/09833245  
GENERAL INFORMATION:  
APPLICANT: Human Genome Sciences, Inc.  
TITLE OF INVENTION: Albumin Fusion Proteins  
FILE REFERENCE: PFS46PCT  
CURRENT APPLICATION NUMBER: US/09/833, 245  
CURRENT FILING DATE: 2001-04-12  
PRIOR APPLICATION NUMBER: 60/229, 358  
PRIOR FILING DATE: 2000-04-12  
PRIOR APPLICATION NUMBER: 60/256, 931  
PRIOR FILING DATE: 2000-12-21  
PRIOR APPLICATION NUMBER: 60/199, 384  
PRIOR FILING DATE: 2000-04-25  
NUMBER OF SEQ ID NOS: 2267



Qy 1 MRYLLPSVLLGTAFTYVLANGVWRLLSAFLPARFYQALDRLCYCVQSWLFFPENYTG 60  
Db 1 MRYLLPSVLLGTAFTYVLANGVWRLLSAFLPARFYQALDRLCYCVQSWLFFPENYTG 60  
Qy 61 VOILLYGLPKNKENIITYLANHSTVDMIVADIATRONALGHVRYVLEKGLKMLPLYGC 120  
Db 61 VOILLYGLPKNKENIITYLANHSTVDMIVADIATRONALGHVRYVLEKGLKMLPLYGC 120  
Qy 121 YFAOHCIGYVRSKAFENKRNKLSQSYVDAGTPEYVIFPEGGRYNPEQTKVLSASQAF 180  
Db 121 YFAOHCIGYVRSKAFENKRNKLSQSYVDAGTPEYVIFPEGGRYNPEQTKVLSASQAF 180  
Qy 181 AAQRCGLAVLKHVLPRIKATHVAFDCMKNYLDAIYDVTVYEGKDDGG 228  
Db 181 AAQRCGLAVLKHVLPRIKATHVAFDCMKNYLDAIYDVTVYEGKDDGG 228

RESULT 12  
US-09-901-484A-4  
: Sequence 4, Application US/09901484A  
: GENERAL INFORMATION:  
: APPLICANT: Cohen, Daniel  
: APPLICANT: Blumenfeld, Marta  
: APPLICANT: Chumakov, Ilya  
: APPLICANT: Bougueleret, Lydie  
: TITLE OF INVENTION: Prostate Cancer Gene  
: FILE REFERENCE: GEN-T11XC3D2  
: CURRENT APPLICATION NUMBER: US/09/901,484A  
: CURRENT FILING DATE: 2001-07-09  
: PRIOR APPLICATION NUMBER: US 08/996,306  
: PRIOR FILING DATE: 1997-12-22  
: PRIOR APPLICATION NUMBER: US 60/099,658  
: PRIOR FILING DATE: 1998-09-09  
: PRIOR APPLICATION NUMBER: US 09/218,207  
: PRIOR FILING DATE: 1998-12-22  
: PRIOR APPLICATION NUMBER: US 09/338,907  
: PRIOR FILING DATE: 1999-06-23  
: PRIOR APPLICATION NUMBER: US 09/853,526  
: PRIOR FILING DATE: 2001-05-11  
: NUMBER OF SEQ ID NOS: 578  
: SOFTWARE: PatentIn version 3.1  
: SEQ ID NO 4  
: LENGTH: 353  
: TYPE: PRT  
: ORGANISM: Homo sapiens  
: FEATURE:  
: NAME/KEY: HELIX  
: LOCATION: (1)..(33)  
: OTHER INFORMATION: Rao and Argos identification method, potential helix  
: NAME/KEY: HELIX  
: LOCATION: (4)..(20)  
: OTHER INFORMATION: Klein, Kanehisa and Delisi identification method, potential helix  
: NAME/KEY: HELIX  
: LOCATION: (4)..(24)  
: OTHER INFORMATION: Eisenberg, Schwarz, Komaromy, Wall identification method,  
: OTHER INFORMATION: potential helix  
: NAME/KEY: LIPID  
: LOCATION: (12)..(16)  
: OTHER INFORMATION: MYRISTATE, Prosite match  
: NAME/KEY: HELIX  
: LOCATION: (50)..(70)  
: OTHER INFORMATION: Eisenberg, Schwarz, Komaromy, Wall identification method,  
: OTHER INFORMATION: potential helix  
: NAME/KEY: CARBOHYD  
: LOCATION: (57)..(59)  
: OTHER INFORMATION: Prosite match  
: NAME/KEY: HELIX  
: LOCATION: (76)..(96)  
: OTHER INFORMATION: Eisenberg, Schwarz, Komaromy, Wall identification method,  
: OTHER INFORMATION: potential helix  
: NAME/KEY: MOD\_RES  
: LOCATION: (78)..(78)

OTHER INFORMATION: PHOSPHORYLATION, potential Tyrosine kinase site, Prosite match  
: NAME/KEY: MOD\_RES  
: LOCATION: (84)..(84)  
: OTHER INFORMATION: PHOSPHORYLATION, potential caseine kinase II site, Prosite mat  
: NAME/KEY: SITE  
: LOCATION: (94)..(115)  
: OTHER INFORMATION: Potential Leucine Zipper site, Prosite match  
: NAME/KEY: LIPID  
: LOCATION: (119)..(123)  
: OTHER INFORMATION: MYRISTATE, Prosite match  
: NAME/KEY: MOD\_RES  
: LOCATION: (133)..(133)  
: OTHER INFORMATION: PHOSPHORYLATION, potential protein kinase C, Prosite match  
: NAME/KEY: MOD\_RES  
: LOCATION: (147)..(147)  
: OTHER INFORMATION: PHOSPHORYLATION, potential caseine kinase II, Prosite match  
: NAME/KEY: MOD\_RES  
: LOCATION: (194)..(194)  
: OTHER INFORMATION: PHOSPHORYLATION, potential Tyrosine kinase site, Prosite match  
: NAME/KEY: MOD\_RES  
: LOCATION: (215)..(215)  
: OTHER INFORMATION: PHOSPHORYLATION, potential protein kinase C, Prosite match  
: NAME/KEY: MOD\_RES  
: LOCATION: (221)..(221)  
: OTHER INFORMATION: SULFATATION, Prosite match  
: NAME/KEY: MOD\_RES  
: LOCATION: (233)..(233)  
: OTHER INFORMATION: PHOSPHORYLATION, potential cAMP and cGMP dependant protein kin  
: NAME/KEY: MOD\_RES  
: LOCATION: (235)..(235)  
: OTHER INFORMATION: PHOSPHORYLATION, potential caseine kinase II site, Prosite mat  
: NAME/KEY: MOD\_RES  
: LOCATION: (306)..(306)  
: OTHER INFORMATION: PHOSPHORYLATION, potential protein kinase C, Prosite match  
: NAME/KEY: HELIX  
: LOCATION: (310)..(330)  
: OTHER INFORMATION: Eisenberg, Schwarz, Komaromy, Wall identification method,  
: OTHER INFORMATION: potential helix  
: NAME/KEY: LIPID  
: LOCATION: (319)..(323)  
: OTHER INFORMATION: MYRISTATE, Prosite match  
: NAME/KEY: LIPID  
: LOCATION: (323)..(327)  
: OTHER INFORMATION: MYRISTATE, Prosite match  
: NAME/KEY: MOD\_RES  
: LOCATION: (329)..(329)  
: OTHER INFORMATION: AMIDATION, Prosite match  
: NAME/KEY: HELIX  
: LOCATION: (333)..(353)  
: OTHER INFORMATION: Eisenberg, Schwarz, Komaromy, Wall identification method,  
: OTHER INFORMATION: potential helix  
: NAME/KEY: LIPID  
: LOCATION: (341)..(345)  
: OTHER INFORMATION: MYRISTATE, Prosite match  
: NAME/KEY: MOD\_RES  
: LOCATION: (350)..(350)  
: OTHER INFORMATION: PHOSPHORYLATION, potential protein kinase C, Prosite match  
US-09-901-484A-4

Query Match 100.0%; Score 1203; DB 23; Length 353;  
Best Local Similarity 100.0%; Pred. No. 1,4e-119;  
Matches 228; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MRYLLPSVLLGTAFTYVLANGVWRLLSAFLPARFYQALDRLCYCVQSWLFFPENYTG 60  
Db 1 MRYLLPSVLLGTAFTYVLANGVWRLLSAFLPARFYQALDRLCYCVQSWLFFPENYTG 60  
Qy 61 VOILLYGLPKNKENIITYLANHSTVDMIVADIATRONALGHVRYVLEKGLKMLPLYGC 120  
Db 61 VOILLYGLPKNKENIITYLANHSTVDMIVADIATRONALGHVRYVLEKGLKMLPLYGC 120

QY	121	YFAOHGGIYYVRSKAFNFKKEKRNKLOSYVAGFPMYLYIPEEGRYNRPBQTKYLSAQAF	160
	121	YFAOHGGIYYVRSKAFNFKKEKRNKLOSYVAGFPMYLYIPEEGRYNRPBQTKYLSAQAF	160
Db	121	YFAOHGGIYYVRSKAFNFKKEKRNKLOSYVAGFPMYLYIPEEGRYNRPBQTKYLSAQAF	160
QY	181	AAOGLAVLKHVLTPIKATHVAFCDCKNKNTLDAIYDVTYVYEGKDDGG	228
Db	181	AAOGLAVLKHVLTPIKATHVAFCDCKNKNTLDAIYDVTYVYEGKDDGG	228

```

RESULT .13
US-09-915-582-59
; Sequence 59, Application US/09915582
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: 17 Human secreted Proteins
; FILE REFERENCE: PS723p1
; CURRENT APPLICATION NUMBER: US/09/915,582
; CURRENT FILING DATE: 2001-07-27
; PRIOR APPLICATION NUMBER: PCT/US01/01431
; PRIOR FILING DATE: 2001-01-17
; PRIOR APPLICATION NUMBER: 60/179,065
; PRIOR FILING DATE: 2000-01-31
; PRIOR APPLICATION NUMBER: 60/180,638
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: 60/231,968
; PRIOR FILING DATE: 2000-09-12
; NUMBER OF SEQ ID NOS: 97
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 59
; LENGTH: 353
; TYPE: PRP
; ORGANISM: Homo sapiens
US-09-915-582-59

```

```

RESULT 14
US-09-915-582-79
Sequence 79, Application US/09915582
GENERAL INFORMATION:
APPLICANT: Rosen et al.
TITLE OF INVENTION: 17 Human Secreted Proteins
FILE REFERENCE: PS723PI
CURRENT APPLICATION NUMBER: US/09/915,582
CURRENT FILING DATE: 2001-07-27
PRIOR APPLICATION NUMBER: PCT/US01/01431
PRIOR FILING DATE: 2001-01-17
PRIOR APPLICATION NUMBER: 60/119,065
PRIOR FILING DATE: 2000-01-31
PRIOR APPLICATION NUMBER: 60/180,628
PRIOR FILING DATE: 2000-02-04
PRIOR APPLICATION NUMBER: 60/231,968
PRIOR FILING DATE: 2000-09-12

```

```

:      NUMBER OF SEQ. ID NOS: 97
:      SOFTWARE: PatentIn Ver. 2.0
:      SEQ ID NO 79
:      LENGTH: 353
:      TYPE: prt
:      ORGANISM: Homo sapiens
:      OS-09-915-582-79

```

```

RESULT      15
US-60-099-658-4
? sequence 4, Application US/60099658
? GENERAL INFORMATION:
? APPLICANT: Cohen, Daniel
? APPLICANT: Chumakov, Ilya
? APPLICANT: Blumenfeld, Marta
? APPLICANT: Bouguenleret, Lydie
? TITLE OF INVENTION: Prostate cancer gene
? NUMBER OF SEQUENCES: 99
? CORRESPONDENCE ADDRESS:
? ADDRESSEE: Knobbe, Martens, Olson & Bear
? STREET: 501 West Broadway
? CITY: San Diego
? STATE: California
? COUNTRY: USA
? ZIP: 92101-3505
? COMPUTER READABLE FORM:
? MEDIUM TYPE: Floppy Disk
? COMPUTER: IBM PC compatible
? OPERATING SYSTEM: WIN95
? SOFTWARE: Word
? CURRENT APPLICATION DATA:
? APPLICATION NUMBER: US/60/099,658
? FILING DATE:
? CLASSIFICATION:
? ATTORNEY/AGENT INFORMATION:
? NAME: Israelisen, Ned A.
? REGISTRATION NUMBER: 29,655
? REFERENCE/DOCKET NUMBER: GENSET. 018APR
? TELECOMMUNICATION INFORMATION:
? TELEPHONE: (619) 235-8550
? TELEFAX: (619) 235-0176
? INFORMATION FOR SEQ ID NO: 4:
? SEQUENCE CHARACTERISTICS:
? LENGTH: 353 amino acids
? TYPE: AMINO ACID
? STRANDEDNESS: SINGLE
? TOPOLOGY: LINEAR
? MOLECULE TYPE: PROTEIN
? ORIGINAL SOURCE:
? ORGANISM: Homo sapiens
? FEATURE:
?
```

NAME/KEY: potential Transmembrane helix  
LOCATION: 1..33  
IDENTIFICATION METHOD: Rao and Argos method  
FEATURE:  
NAME/KEY: potential Transmembrane helix  
LOCATION: 4..20  
IDENTIFICATION METHOD: Klein, Kanehisa and Delisi method  
FEATURE:  
NAME/KEY: potential Transmembrane helix  
LOCATION: 4..24  
IDENTIFICATION METHOD: Eisenberg, Schwarz, Komarony  
IDENTIFICATION METHOD: and Wall method  
FEATURE:  
NAME/KEY: potential N-myristoylation site  
LOCATION: 12  
IDENTIFICATION METHOD: prosite match  
FEATURE:  
NAME/KEY: potential Transmembrane helix  
LOCATION: 50..70  
IDENTIFICATION METHOD: Eisenberg, Schwarz, Komarony and Wall method  
FEATURE:  
NAME/KEY: potential N-glycosylation site  
LOCATION: 57  
IDENTIFICATION METHOD: prosite match  
FEATURE:  
NAME/KEY: potential Transmembrane helix  
LOCATION: 76..96  
IDENTIFICATION METHOD: Eisenberg, Schwarz, Komarony and Wall method  
FEATURE:  
NAME/KEY: potential Tyrosine kinase phosphorylation site  
LOCATION: 78  
IDENTIFICATION METHOD: prosite match  
FEATURE:  
NAME/KEY: potential Caseln kinase II phosphorylation site  
LOCATION: 84  
IDENTIFICATION METHOD: prosite match  
FEATURE:  
NAME/KEY: potential Leucine zipper pattern  
LOCATION: 94..115  
IDENTIFICATION METHOD: prosite match  
FEATURE:  
NAME/KEY: potential N-myristoylation site  
LOCATION: 119  
IDENTIFICATION METHOD: prosite match  
FEATURE:  
NAME/KEY: potential Protein kinase C phosphorylation site  
LOCATION: 133  
IDENTIFICATION METHOD: prosite match  
FEATURE:  
NAME/KEY: potential Caseln kinase II phosphorylation site  
LOCATION: 147  
IDENTIFICATION METHOD: prosite match  
FEATURE:  
NAME/KEY: potential Protein kinase C phosphorylation site  
LOCATION: 194  
IDENTIFICATION METHOD: prosite match  
FEATURE:  
NAME/KEY: potential Tyrosine kinase phosphorylation site  
LOCATION: 215  
IDENTIFICATION METHOD: prosite match  
FEATURE:  
NAME/KEY: potential Tyrosine sulfatation site  
LOCATION: 221  
IDENTIFICATION METHOD: prosite match  
FEATURE:  
NAME/KEY: potential cAMP- and cGMP-dependent protein kinase phosphorylation site  
LOCATION: 233  
IDENTIFICATION METHOD: prosite match  
FEATURE:  
NAME/KEY: potential Caseln kinase II phosphorylation site  
LOCATION: 235  
IDENTIFICATION METHOD: prosite match  
FEATURE:

NAME/KEY: potential Protein kinase C phosphorylation site  
LOCATION: 306  
IDENTIFICATION METHOD: prosite match  
FEATURE:  
NAME/KEY: potential Transmembrane helix  
LOCATION: 310..330  
IDENTIFICATION METHOD: Eisenberg, Schwarz, Komarony and Wall method  
FEATURE:  
NAME/KEY: potential N-myristoylation site  
LOCATION: 319  
IDENTIFICATION METHOD: prosite match  
FEATURE:  
NAME/KEY: potential N-myristoylation site  
LOCATION: 323  
IDENTIFICATION METHOD: prosite match  
FEATURE:  
NAME/KEY: potential Amidation site  
LOCATION: 329  
IDENTIFICATION METHOD: prosite match  
FEATURE:  
NAME/KEY: potential Transmembrane helix  
LOCATION: 333..353  
IDENTIFICATION METHOD: Eisenberg, Schwarz, Komarony and Wall method  
FEATURE:  
NAME/KEY: potential N-myristoylation site  
LOCATION: 341  
IDENTIFICATION METHOD: prosite match  
FEATURE:  
NAME/KEY: potential Protein kinase C phosphorylation site  
LOCATION: 350  
IDENTIFICATION METHOD: prosite match  
US-60-099-658-4

Query Match 100.0%; Score 1203; DB 26; Length 353;  
Best Local Similarity 100.0%; Pred. No. 1.4e-119;  
Matches 228; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 MRYLLPSVVLGTAPTYVLAMGWRLSAPLPARYOALDRILCYQSMVLEFFENYTG 60  
Db 1 MRYLLPSVVLGTAPTYVLAMGWRLSAPLPARYOALDRILCYQSMVLEFFENYTG 60  
Qy 61 VQILLYGDLPKRNENIYLANHOSYVDIADILAIQNALGHVRYLKGKMLPLXGC 120  
Db 61 VQILLYGDLPKRNENIYLANHOSYVDIADILAIQNALGHVRYLKGKMLPLXGC 120  
Qy 121 YFAOHGSIYVRSKAKFNEKEKRNKLOSVDAGTPMYLVIFPEGRYVPEOTKYLASQAF 180  
Db 121 YFAOHGSIYVRSKAKFNEKEKRNKLOSVDAGTPMYLVIFPEGRYVPEOTKYLASQAF 180  
Qy 181 AAOGLAVLKHVLPRIKATHVAFDCMKNYLDAIYDVTVVEGKDDG 228  
Db 181 AAOGLAVLKHVLPRIKATHVAFDCMKNYLDAIYDVTVVEGKDDG 228

Search completed: August 28, 2002, 11:14:11  
Job time: 490 sec



---

**THIS PAGE BLANK (USPTO)**

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compen Ltd.

OM protein - protein search, using sw model

Run on: August 28, 2002, 11:15:42 ; Search time 94.08 Seconds  
(without alignments)  
656.370 Million cell updates/sec

Title: US-09-853-526-70  
Perfect score: 1203  
Sequence: 1 MRYLPSVVLGTAPFYVLA.....NYDAIVDFVVEGKDDG 228

Scoring table: BIOSUM62  
Gapop 10.0, Gapext 0.5

Searched: 747981 seqs, 242050750 residues

Total number of hits satisfying chosen parameters: 747981

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Pending\_Patents\_AA\_New.\*  
1: /cgn2\_6/ptodata/1/paa/PCT\_NEW\_COMB.pep.\*  
2: /cgn2\_6/ptodata/1/paa/US06\_NEW\_COMB.pep.\*  
3: /cgn2\_6/ptodata/1/paa/US07\_NEW\_COMB.pep.\*  
4: /cgn2\_6/ptodata/1/paa/US08\_NEW\_COMB.pep.\*  
5: /cgn2\_6/ptodata/1/paa/US09\_NEW\_COMB.pep.\*  
6: /cgn2\_6/ptodata/1/paa/US10\_NEW\_COMB.pep.\*  
7: /cgn2\_6/ptodata/1/paa/US60\_NEW\_COMB.pep.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1203	100.0	364	6	US-10-184-648-63
2	1200	99.8	353	5	US-09-629-469A-13028
3	759	63.1	269	6	US-10-074-045-47
4	291	24.2	375	5	US-09-935-625-11061
5	291	24.2	375	5	US-09-935-625-14197
6	291	24.2	375	5	US-09-935-625-14201
7	291	24.2	375	5	US-09-935-625-23009
8	291	24.2	375	5	US-09-935-625-23025
9	291	24.2	375	5	US-09-935-625-11062
10	288	23.9	351	5	US-09-935-625-14198
11	288	23.9	351	5	US-09-935-625-14202
12	288	23.9	351	5	US-09-935-625-14255
13	288	23.9	351	5	US-09-935-625-23010
14	288	23.9	351	5	US-09-935-625-23026
15	288	23.9	351	5	US-09-935-625-11471
16	288	23.9	351	5	US-09-935-625-11472
17	285	23.7	351	5	US-09-935-625-11472
18	285	23.4	400	7	US-60-391-781-1157
19	275	22.9	343	5	US-09-935-625-11063
20	275	22.9	343	5	US-09-935-625-11473
21	275	22.9	343	5	US-09-935-625-14199
22	275	22.9	343	5	US-09-935-625-14203
23	275	22.9	343	5	US-09-935-625-14256
24	275	22.9	343	5	US-09-935-625-23011
25	275	22.9	343	5	US-09-935-625-23027
26	233	19.4	414	6	US-10-121-062-314

27	233	19.4	414	6	US-10-063-502-102	Sequence 102, App
28	233	19.4	414	6	US-10-063-510-102	Sequence 102, App
29	233	19.4	414	6	US-10-063-512-102	Sequence 102, App
30	233	19.4	414	6	US-10-063-513-102	Sequence 102, App
31	233	19.4	414	6	US-10-063-514-102	Sequence 102, App
32	233	19.4	414	6	US-10-063-515-102	Sequence 102, App
33	233	19.4	414	6	US-10-063-516-102	Sequence 102, App
34	233	19.4	414	6	US-10-063-517-102	Sequence 102, App
35	233	19.4	414	6	US-10-063-518-102	Sequence 102, App
36	233	19.4	414	6	US-10-063-519-102	Sequence 102, App
37	233	19.4	414	6	US-10-063-520-102	Sequence 102, App
38	233	19.4	414	6	US-10-063-521-102	Sequence 102, App
39	233	19.4	414	6	US-10-063-523-102	Sequence 102, App
40	233	19.4	414	6	US-10-063-524-102	Sequence 102, App
41	233	19.4	414	6	US-10-063-525-102	Sequence 102, App
42	233	19.4	414	6	US-10-063-526-102	Sequence 102, App
43	233	19.4	414	6	US-10-063-527-102	Sequence 102, App
44	233	19.4	414	6	US-10-063-528-102	Sequence 102, App
45	233	19.4	414	6	US-10-063-529-102	Sequence 102, App

## ALIGNMENTS

RESULT 1  
US-10-184-648-63  
Sequence 63, Application US/10184648  
GENERAL INFORMATION:  
APPLICANT: Meyers, Rachel E.  
APPLICANT: Williamson, Mark  
APPLICANT: Tsai, Feng-Ying  
APPLICANT: Hunter, John J.  
APPLICANT: Macbeth, Kyle J.  
APPLICANT: Rudolph-Owen, Laura A.  
APPLICANT: Leibly, Kevin R.  
APPLICANT: Kapeller-Libermann, Rosana  
APPLICANT: Olandt, Peter J.  
TITLE OF INVENTION: NOVEL HUMAN TRANSFERASE FAMILY MEMBERS AND USES THEREOF  
FILE REFERENCE: 10448-192001  
CURRENT APPLICATION NUMBER: US/10/184, 648  
CURRENT FILING DATE: 2002-06-27  
PRIOR APPLICATION NUMBER: US 09/815, 028  
PRIOR FILING DATE: 2001-03-22  
PRIOR APPLICATION NUMBER: PCT/US01/09358  
PRIOR FILING DATE: 2001-03-22  
PRIOR APPLICATION NUMBER: US 60/191, 964  
PRIOR FILING DATE: 2000-03-24  
PRIOR APPLICATION NUMBER: US 09/801, 220  
PRIOR FILING DATE: 2001-03-07  
PRIOR APPLICATION NUMBER: PCT/US01/07269  
PRIOR FILING DATE: 2001-03-07  
PRIOR APPLICATION NUMBER: US 60/187, 456  
PRIOR FILING DATE: 2000-03-07  
PRIOR APPLICATION NUMBER: US 09/816, 714  
PRIOR FILING DATE: 2001-03-23  
PRIOR APPLICATION NUMBER: PCT/US01/09468  
PRIOR FILING DATE: 2001-03-23  
PRIOR APPLICATION NUMBER: US 60/191, 865  
PRIOR FILING DATE: 2000-03-24  
PRIOR APPLICATION NUMBER: US 09/844, 948  
PRIOR FILING DATE: 2001-04-27  
PRIOR APPLICATION NUMBER: PCT/US01/13805  
PRIOR FILING DATE: 2001-04-27  
PRIOR APPLICATION NUMBER: US 60/200, 604  
PRIOR FILING DATE: 2000-04-28  
PRIOR APPLICATION NUMBER: US 09/861, 164  
PRIOR FILING DATE: 2001-05-18  
PRIOR APPLICATION NUMBER: PCT/US01/16292  
PRIOR FILING DATE: 2001-05-18  
PRIOR APPLICATION NUMBER: US 60/205, 408  
PRIOR FILING DATE: 2000-05-19  
PRIOR APPLICATION NUMBER: US 09/883, 060  
PRIOR FILING DATE: 2001-06-15

```

; PRIOR APPLICATION NUMBER: PCT/US01/19138
; PRIOR FILING DATE: 2001-06-15
; PRIOR APPLICATION NUMBER: US 60/212,079
; PRIOR FILING DATE: 2000-06-15
; PRIOR APPLICATION NUMBER: US 09/962,678
; PRIOR FILING DATE: 2001-09-25
; PRIOR APPLICATION NUMBER: PCT/US01/29963
; PRIOR FILING DATE: 2001-09-25
; PRIOR APPLICATION NUMBER: US 60/235,044
; PRIOR FILING DATE: 2000-09-25
; PRIOR APPLICATION NUMBER: US 09/973,457
; PRIOR FILING DATE: 2001-10-09
; PRIOR APPLICATION NUMBER: US 60/238,849
; PRIOR FILING DATE: 2000-10-06
; PRIOR APPLICATION NUMBER: US 10/072,285
; PRIOR FILING DATE: 2002-02-08
; PRIOR APPLICATION NUMBER: PCT/US02/03736
; PRIOR FILING DATE: 2002-02-08
; PRIOR APPLICATION NUMBER: US 60/267,494
; PRIOR FILING DATE: 2001-02-08
; PRIOR APPLICATION NUMBER: US 09/817,910
; PRIOR FILING DATE: 2001-03-26
; PRIOR APPLICATION NUMBER: PCT/US01/09633
; PRIOR FILING DATE: 2001-03-26
; PRIOR APPLICATION NUMBER: US 60/192,092
; PRIOR FILING DATE: 2000-03-24
; PRIOR APPLICATION NUMBER: US 09/842,528
; PRIOR FILING DATE: 2001-04-25
; PRIOR APPLICATION NUMBER: PCT/US01/40607
; PRIOR FILING DATE: 2001-04-25
; PRIOR APPLICATION NUMBER: US 60/199,500
; PRIOR FILING DATE: 2000-04-25
; PRIOR APPLICATION NUMBER: US 09/882,836
; PRIOR FILING DATE: 2001-06-15
; PRIOR APPLICATION NUMBER: PCT/US01/19543
; PRIOR FILING DATE: 2001-06-15
; PRIOR APPLICATION NUMBER: US 60/211,730
; PRIOR FILING DATE: 2000-06-15
; PRIOR APPLICATION NUMBER: US 09/882,872
; PRIOR FILING DATE: 2001-06-15
; PRIOR APPLICATION NUMBER: PCT/US01/19153
; PRIOR FILING DATE: 2001-06-15
; PRIOR APPLICATION NUMBER: US 60/212,077
; PRIOR FILING DATE: 2000-06-15
; NUMBER OF SEQ ID NOS: 90
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 63
; LENGTH: 364
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-184-648-63
```

Query Match 100.0%; Score 1203; DB 6; Length 364;  
Best Local Similarity 100.0%; Pred. No. 4,3e-116;  
Matches 228; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 MRYLLPSVLLGTAPTYVLANGVWRLLSAFLPARFYQALDRLKCVQSMVLPFFENYTG 60
    |||||||
DB 12 MRYLLPSVLLGTAPTYVLANGVWRLLSAFLPARFYQALDRLKCVQSMVLPFFENYTG 71
    |||||||
QY 61 VOILLYGDLPKKNENIYLANHSTVDIYADILAIQONALGHVRYVLEKGLKMLPLYGC 120
    |||||||
DB 72 VOILLYGDLPKKNENIYLANHSTVDIYADILAIQONALGHVRYVLEKGLKMLPLYGC 131
    |||||||
QY 121 YFAOHGIIYVRSKAKFNEKEKRNKLOSVDAGTPLYVIEPEGTRYNPEOTKVLASQAF 180
    |||||||
DB 132 YFAOHGIIYVRSKAKFNEKEKRNKLOSVDAGTPLYVIEPEGTRYNPEOTKVLASQAF 191
    |||||||
QY 181 AAQGLAVLKHVLTPIRIKATHVAFDCKKNYIDATYDVTYVYEGKDDG 228
    |||||||
DB 192 AAQGLAVLKHVLTPIRIKATHVAFDCKKNYIDATYDVTYVYEGKDDG 239
    |||||||
```

```

RESULT 2
US-09-629-469A-13028
; Sequence 13028, Application US/09629469A
; GENERAL INFORMATION:
; APPLICANT: OTA, TOSHIO
; APPLICANT: ISOGAI, TAKAO
; APPLICANT: NISHIKAWA, TETSUO
; APPLICANT: HAYASHI, KOJI
; APPLICANT: SAITO, KAORU
; APPLICANT: YAMAMOTO, JUNICHI
; APPLICANT: ISHII, SHIZUKO
; APPLICANT: SUGITAMA, TOMOYASU
; APPLICANT: WAKAMATSU, AI
; APPLICANT: MAGAI, KEIICHI
; APPLICANT: OTSUKI, TETSUJI
; TITLE OF INVENTION: PRIMERS FOR SYNTHESIZING FULL-LENGTH CDNA AND THEIR USE
; FILE REFERENCE: 084335/0123
; CURRENT FILING DATE: 2000-07-28
; PRIOR APPLICATION NUMBER: US/09/629,469A
; PRIOR FILING DATE: 1999-07-29
; PRIOR APPLICATION NUMBER: JP 1999-248036
; PRIOR FILING DATE: 1999-08-27
; PRIOR APPLICATION NUMBER: JP 2000-118776
; PRIOR FILING DATE: 2000-01-11
; PRIOR APPLICATION NUMBER: JP 2000-183767
; PRIOR FILING DATE: 2000-05-02
; PRIOR APPLICATION NUMBER: JP 2000-241899
; PRIOR FILING DATE: 2000-06-09
; PRIOR APPLICATION NUMBER: 60/159,590
; PRIOR FILING DATE: 1999-10-18
; PRIOR APPLICATION NUMBER: 60/183,322
; PRIOR FILING DATE: 2000-02-17
; NUMBER OF SEQ ID NOS: 19025
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 13028
; LENGTH: 353
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-629-469A-13028
```

Query Match 99.8%; Score 1200; DB 5; Length 353;  
Best Local Similarity 99.6%; Pred. No. 8,4e-116;  
Matches 227; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 MRYLLPSVLLGTAPTYVLANGVWRLLSAFLPARFYQALDRLKCVQSMVLPFFENYTG 60
    |||||||
DB 1 MRYLLPSVLLGTAPTYVLANGVWRLLSAFLPARFYQALDRLKCVQSMVLPFFENYTG 60
    |||||||
QY 61 VOILLYGDLPKKNENIYLANHSTVDIYADILAIQONALGHVRYVLEKGLKMLPLYGC 120
    |||||||
DB 61 VOILLYGDLPKKNENIYLANHSTVDIYADILAIQONALGHVRYVLEKGLKMLPLYGC 120
    |||||||
QY 121 YFAOHGIIYVRSKAKFNEKEKRNKLOSVDAGTPLYVIEPEGTRYNPEOTKVLASQAF 180
    |||||||
DB 121 YFAOHGIIYVRSKAKFNEKEKRNKLOSVDAGTPLYVIEPEGTRYNPEOTKVLASQAF 180
    |||||||
QY 181 AAQGLAVLKHVLTPIRIKATHVAFDCKKNYIDATYDVTYVYEGKDDG 228
    |||||||
DB 181 AAQGLAVLKHVLTPIRIKATHVAFDCKKNYIDATYDVTYVYEGKDDG 228
    |||||||
```

RESULT 3  
US-10-074-045-47  
; Sequence 47, Application US/10074045  
; GENERAL INFORMATION:  
; APPLICANT: Rosen et al.  
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies  
; FILE REFERENCE: PT221C1  
; CURRENT APPLICATION NUMBER: US/10/074,045  
; CURRENT FILING DATE: 2002-02-14

```
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 74
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 47
; LENGTH: 269
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-074-045-47
```

```
Query Match          63.1%; Score 759; DB 6; Length 269;
Best Local Similarity 100.0%; Pred. No. 2.7e-70;
Matches 144; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY 85 TWDVIADILAIROMALGHVRYLKEGLKMLPLGYGCPAHOHGIIYVRSKAFNEKEMRNK 144
DB 1 TYDVIADILAIROMALGHVRYLKEGLKMLPLGYGCPAHOHGIIYVRSKAFNEKEMRNK 60
OY 145 LOSYVDAGTPMYLTFPEEGTRYNPEOTKVLASQAFAOAGLAVLKHLVLPRIKATHVAF 204
DB 61 LOSYVDAGTPMYLTFPEEGTRYNPEOTKVLASQAFAOAGLAVLKHLVLPRIKATHVAF 120
OY 205 DCMKNTLDAIYVTVYVEKDDGG 228
DB 121 DCMKNTLDAIYVTVYVEKDDGG 144
```

## RESULT 4

```
US-09-935-625-11061
; Sequence 11061, Application US/09935625
; GENERAL INFORMATION:
; APPLICANT: N. ALEXANDROV et al.
; TITLE OF INVENTION: POLYNUCLEOTIDES, POLYPEPTIDES, CELLS, AND METHODS THEREOF CAPABLE
; FILE REFERENCE: 2750-1481P
; CURRENT APPLICATION NUMBER: US/09/935,625
; CURRENT FILING DATE: 2001-08-24
; NUMBER OF SEQ ID NOS: 33136
; SEQ ID NO 11061
; LENGTH: 375
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; NAME/KEY: peptide
; LOCATION: 1..375
; OTHER INFORMATION: Ceres Seq. ID no. 1394370
US-09-935-625-11061
```

```
Query Match          24.2%; Score 291; DB 5; Length 375;
Best Local Similarity 31.3%; Pred. No. 1.2e-21;
Matches 71; Conservative 50; Mismatches 94; Indels 12; Gaps 3;
```

```
OY 1 MRYLPSVVLGTAPTYVILAMG-----VWRLSAFLPARFYQALDRLXCVYOSMVLFFP 55
DB 18 LRGIIICMLVSTAFMMLIFMGFLSAVILRLFS---IRYSRCVSPFFGSMALMPFLF 73
OY 56 ENYTGVOILLYGDLPKNKENIITLANHOSTVDVIADILAIROMALGHVRYLKEGLKWL 115
DB 74 EKINKTKVIFSGDKVPCEDRVLLIANHRTVDMMYFMDLALRRGOIGNIKYVLSIMKL 133
OY 116 PLVGYFAOHGIIYVRSKAFNEKEMRNKLOSVDAGTPMYLTFPEEGTRYNPEOTKVL 175
DB 134 PLFGMAFHLEFIPVERRMEVDKANLROIYSSFKDRDALMLALFPEGDTYTAKCO--- 190
OY 176 ASQAFAOAGLAVLKHLVLPRIKATHVAFDCMKNYDAIYVTVYVE 222
DB 191 RSKKFAENGSLPLNNVLLPRTKGFVSCIQELSCSIDAVYDVITGYK 237
```

## RESULT 5

```
US-09-935-625-14197
; Sequence 14197, Application US/09935625
```

```
; GENERAL INFORMATION:
; APPLICANT: N. ALEXANDROV et al.
; TITLE OF INVENTION: POLYNUCLEOTIDES, POLYPEPTIDES, CELLS, AND METHODS THEREOF CAPA
; FILE REFERENCE: 2750-1481P
; CURRENT APPLICATION NUMBER: US/09/935,625
; CURRENT FILING DATE: 2001-08-24
; NUMBER OF SEQ ID NOS: 33136
; SEQ ID NO 14197
; LENGTH: 375
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; NAME/KEY: peptide
; LOCATION: 1..375
; OTHER INFORMATION: Ceres Seq. ID no. 3441314
US-09-935-625-14197
```

```
Query Match          24.2%; Score 291; DB 5; Length 375;
Best Local Similarity 31.3%; Pred. No. 1.2e-21;
Matches 71; Conservative 50; Mismatches 94; Indels 12; Gaps 3;
```

```
OY 1 MRYLPSVVLGTAPTYVILAMG-----VWRLSAFLPARFYQALDRLXCVYOSMVLFFP 55
DB 18 LRGIIICMLVSTAFMMLIFMGFLSAVILRLFS---IRYSRCVSPFFGSMALMPFLF 73
OY 56 ENYTGVOILLYGDLPKNKENIITLANHOSTVDVIADILAIROMALGHVRYLKEGLKWL 115
DB 74 EKINKTKVIFSGDKVPCEDRVLLIANHRTVDMMYFMDLALRRGOIGNIKYVLSIMKL 133
OY 116 PLVGYFAOHGIIYVRSKAFNEKEMRNKLOSVDAGTPMYLTFPEEGTRYNPEOTKVL 175
DB 134 PLFGMAFHLEFIPVERRMEVDKANLROIYSSFKDRDALMLALFPEGDTYTAKCO--- 190
OY 176 ASQAFAOAGLAVLKHLVLPRIKATHVAFDCMKNYDAIYVTVYVE 222
DB 191 RSKKFAENGSLPLNNVLLPRTKGFVSCIQELSCSIDAVYDVITGYK 237
```

## RESULT 6

```
US-09-935-625-14201
; Sequence 14201, Application US/09935625
; GENERAL INFORMATION:
; APPLICANT: N. ALEXANDROV et al.
; TITLE OF INVENTION: POLYNUCLEOTIDES, POLYPEPTIDES, CELLS, AND METHODS THEREOF CAPA
; FILE REFERENCE: 2750-1481P
; CURRENT APPLICATION NUMBER: US/09/935,625
; CURRENT FILING DATE: 2001-08-24
; NUMBER OF SEQ ID NOS: 33136
; SEQ ID NO 14201
; LENGTH: 375
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; NAME/KEY: peptide
; LOCATION: 1..375
; OTHER INFORMATION: Ceres Seq. ID no. 3441318
US-09-935-625-14201
```

```
Query Match          24.2%; Score 291; DB 5; Length 375;
Best Local Similarity 31.3%; Pred. No. 1.2e-21;
Matches 71; Conservative 50; Mismatches 94; Indels 12; Gaps 3;
```

```
OY 1 MRYLPSVVLGTAPTYVILAMG-----VWRLSAFLPARFYQALDRLXCVYOSMVLFFP 55
DB 18 LRGIIICMLVSTAFMMLIFMGFLSAVILRLFS---IRYSRCVSPFFGSMALMPFLF 73
OY 56 ENYTGVOILLYGDLPKNKENIITLANHOSTVDVIADILAIROMALGHVRYLKEGLKWL 115
DB 74 EKINKTKVIFSGDKVPCEDRVLLIANHRTVDMMYFMDLALRRGOIGNIKYVLSIMKL 133
```

OY 116 PLVGCYFAOHGIIYKRSKAFNEKEMRNKLSQSYVDAGTPLYLVEFGSTRYNPEOTKVL 175  
Db 134 PLFGMAHLEFEFIVERKWEDEANLQIYSSFKDPRDALMLALFPGTDYTEAKCQ--- 190  
OY 176 ASQAFPAORGLAVLKHVLTPIKATHVAFDCMKNYLDAIYDVYVE 222  
Db 191 RSKKFAENGLPIILNNVLLPRTKGFVSCLOELSCSLDAVYDVITGYK 237

## RESULT 7

US-09-935-625-14254  
; Sequence 14254, Application US/09935625  
; GENERAL INFORMATION:  
; APPLICANT: N. ALEXANDROV et al.  
; TITLE OF INVENTION: POLYNUCLEOTIDES, POLYPEPTIDES, CELLS, AND METHODS THEREOF CAPABLE  
; FILE REFERENCE: 2750-1481P  
; CURRENT APPLICATION NUMBER: US/09/935,625  
; NUMBER OF SEQ ID NOS: 2001-08-24  
; SEQ ID NO 14254  
; LENGTH: 375  
; TYPE: PRP  
; ORGANISM: Arabidopsis thaliana  
; FEATURE:  
; NAME/KEY: Peptide  
; LOCATION: 1..375  
; OTHER INFORMATION: Ceres Seq. ID no. 3443248  
US-09-935-625-14254

## Query Match

Best Local Similarity 24.2%; Score 291; DB 5; Length 375;  
Matches 71; Conservative 50; Mismatches 94; Indels 12; Gaps 3;

OY 1 MRYLLPSVLLGTAPTYVLAMG-----VWRLLSAFLPAREYQALDRLYCYQSMVLEFF 55  
Db 18 LRGIICLMVIVSTAFMMLIFWGFSAVVLRLFS---IRYSRKCVSFFGSMALMPFLF 73  
OY 56 ENTGVOIILLYGDLPRKKNENIYLANHOSYVDWIVADILAIROMALGHVRYLKEGLKWL 115  
Db 74 EKINKTVITSGDKVPCEDRVLLIANHRTEDWMEFDLAIKQIGINIKIVLSSIMKL 133  
OY 116 PLVGCYFAOHGIIYKRSKAFNEKEMRNKLSQSYVDAGTPLYLVEFGSTRYNPEOTKVL 175  
Db 134 PLFGMAHLEFEFIVERKWEDEANLQIYSSFKDPRDALMLALFPGTDYTEAKCQ--- 190  
OY 176 ASQAFPAORGLAVLKHVLTPIKATHVAFDCMKNYLDAIYDVYVE 222  
Db 191 RSKKFAENGLPIILNNVLLPRTKGFVSCLOELSCSLDAVYDVITGYK 237

## RESULT 8

US-09-935-625-23009  
; Sequence 23009, Application US/09935625  
; GENERAL INFORMATION:  
; APPLICANT: N. ALEXANDROV et al.  
; TITLE OF INVENTION: POLYNUCLEOTIDES, POLYPEPTIDES, CELLS, AND METHODS THEREOF CAPABLE  
; FILE REFERENCE: 2750-1481P  
; CURRENT APPLICATION NUMBER: US/09/935,625  
; NUMBER OF SEQ ID NOS: 2001-08-24  
; SEQ ID NO 23009  
; LENGTH: 375  
; TYPE: PRP  
; ORGANISM: Arabidopsis thaliana  
; FEATURE:  
; NAME/KEY: Peptide  
; LOCATION: 1..375  
; OTHER INFORMATION: Ceres Seq. ID no. 3441318  
US-09-935-625-23009

Query Match 24.2%; Score 291; DB 5; Length 375;  
Best Local Similarity 31.3%; Pred. No. 1.2e-21;  
Matches 71; Conservative 50; Mismatches 94; Indels 12; Gaps 3;

OY 1 MRYLLPSVLLGTAPTYVLAMG-----VWRLLSAFLPAREYQALDRLYCYQSMVLEFF 55  
Db 18 LRGIICLMVIVSTAFMMLIFWGFSAVVLRLFS---IRYSRKCVSFFGSMALMPFLF 73  
OY 56 ENTGVOIILLYGDLPRKKNENIYLANHOSYVDWIVADILAIROMALGHVRYLKEGLKWL 115  
Db 74 EKINKTVITSGDKVPCEDRVLLIANHRTEDWMEFDLAIKQIGINIKIVLSSIMKL 133  
OY 116 PLVGCYFAOHGIIYKRSKAFNEKEMRNKLSQSYVDAGTPLYLVEFGSTRYNPEOTKVL 175  
Db 134 PLFGMAHLEFEFIVERKWEDEANLQIYSSFKDPRDALMLALFPGTDYTEAKCQ--- 190  
OY 176 ASQAFPAORGLAVLKHVLTPIKATHVAFDCMKNYLDAIYDVYVE 222  
Db 191 RSKKFAENGLPIILNNVLLPRTKGFVSCLOELSCSLDAVYDVITGYK 237

## RESULT 9

US-09-935-625-23025  
; Sequence 23025, Application US/09935625  
; GENERAL INFORMATION:  
; APPLICANT: N. ALEXANDROV et al.  
; TITLE OF INVENTION: POLYNUCLEOTIDES, POLYPEPTIDES, CELLS, AND METHODS THEREOF CAPA  
; FILE REFERENCE: 2750-1481P  
; CURRENT APPLICATION NUMBER: US/09/935,625  
; NUMBER OF SEQ ID NOS: 2001-08-24  
; SEQ ID NO 23025  
; LENGTH: 375  
; TYPE: PRP  
; ORGANISM: Arabidopsis thaliana  
; FEATURE:  
; NAME/KEY: Peptide  
; LOCATION: 1..375  
; OTHER INFORMATION: Ceres Seq. ID no. 3443248  
US-09-935-625-23025

Query Match 24.2%; Score 291; DB 5; Length 375;  
Best Local Similarity 31.3%; Pred. No. 1.2e-21;  
Matches 71; Conservative 50; Mismatches 94; Indels 12; Gaps 3;

OY 1 MRYLLPSVLLGTAPTYVLAMG-----VWRLLSAFLPAREYQALDRLYCYQSMVLEFF 55  
Db 18 LRGIICLMVIVSTAFMMLIFWGFSAVVLRLFS---IRYSRKCVSFFGSMALMPFLF 73  
OY 56 ENTGVOIILLYGDLPRKKNENIYLANHOSYVDWIVADILAIROMALGHVRYLKEGLKWL 115  
Db 74 EKINKTVITSGDKVPCEDRVLLIANHRTEDWMEFDLAIKQIGINIKIVLSSIMKL 133  
OY 116 PLVGCYFAOHGIIYKRSKAFNEKEMRNKLSQSYVDAGTPLYLVEFGSTRYNPEOTKVL 175  
Db 134 PLFGMAHLEFEFIVERKWEDEANLQIYSSFKDPRDALMLALFPGTDYTEAKCQ--- 190  
OY 176 ASQAFPAORGLAVLKHVLTPIKATHVAFDCMKNYLDAIYDVYVE 222  
Db 191 RSKKFAENGLPIILNNVLLPRTKGFVSCLOELSCSLDAVYDVITGYK 237

## RESULT 10

US-09-935-625-11062  
; Sequence 11062, Application US/09935625  
; GENERAL INFORMATION:  
; APPLICANT: N. ALEXANDROV et al.  
; TITLE OF INVENTION: POLYNUCLEOTIDES, POLYPEPTIDES, CELLS, AND METHODS THEREOF CAPA  
; FILE REFERENCE: 2750-1481P  
; CURRENT APPLICATION NUMBER: US/09/935,625  
; NUMBER OF SEQ ID NOS: 2001-08-24  
; SEQ ID NO 11062  
; LENGTH: 375  
; TYPE: PRP  
; ORGANISM: Arabidopsis thaliana  
; FEATURE:  
; NAME/KEY: Peptide  
; LOCATION: 1..375  
; OTHER INFORMATION: Ceres Seq. ID no. 3441318  
US-09-935-625-11062

```
FILE REFERENCE: 2750-1481P
CURRENT APPLICATION NUMBER: US/09/935,625
CURRENT FILING DATE: 2001-08-24
NUMBER OF SEQ ID NOS: 33136
SEQ ID NO 11062
LENGTH: 351
TYPE: PRP
ORGANISM: Arabidopsis thaliana
FEATURE:
NAME/KEY: peptide
LOCATION: 1..351
OTHER INFORMATION: Ceres Seq. ID no. 1394371
US-09-935-625-11062
```

```
Query Match
Best Local Similarity 23.9%; Score 288; DB 5; Length 351;
Matches 70; Conservative 47; Mismatches 91; Indels 12; Gaps 3;
```

```
QY 8 VVLTGAPTYVLAAG-----VWRLSAFLPARFYQALDRLCYVOSMVLFFPENTGYQ 62
DB 1 MVLVSTAFMMLIFWGLSAVLRIFS---IRSRKCVSFFSGSWALMPLEKINKTK 56
QY 63 ILIYGLPKKNENIYLANHSTVDWIVADILAIQNALGHVRYVLEKGLKMLPLYGCYF 122
DB 57 VIFSGDKVPCEDRVLLIANHRTVDWVYFMDLALRKQIGNIKYVLSLMKLPFGWAF 116
QY 123 AOHGGLYVRSKAFENKEMKNLQSYVDAGTPMYLVIFPEGTRYNPEQTKVLSAQAFAA 182
DB 117 HLEFEFIVERMVEDEANLRQIVSSFKDPDRLMLALFPEGTDYTEAKCO---RSKKFAA 173
QY 183 ORGLAVLKHYLTPRIKATHVAFDCMKNYLDIAYDVTVYE 222
DB 174 ENGLPILNNVLLPRTKGVSCLQELSCSDAVYDVITIGYK 213
```

```
RESULT 11
US-09-935-625-14198
; Sequence 14198, Application US/09935625
; GENERAL INFORMATION:
; APPLICANT: N. ALEXANDROV et al.
; TITLE OF INVENTION: POLYNUCLEOTIDES, POLYPEPTIDES, CELLS, AND METHODS THEREOF CAPABLE
; FILE REFERENCE: 2750-1481P
; CURRENT APPLICATION NUMBER: US/09/935,625
; CURRENT FILING DATE: 2001-08-24
; NUMBER OF SEQ ID NOS: 33136
; SEQ ID NO 14198
; LENGTH: 351
; TYPE: PRP
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; NAME/KEY: peptide
; LOCATION: 1..351
; OTHER INFORMATION: Ceres Seq. ID no. 3441315
US-09-935-625-14198
```

```
Query Match
Best Local Similarity 23.9%; Score 288; DB 5; Length 351;
Matches 70; Conservative 47; Mismatches 91; Indels 12; Gaps 3;
```

```
QY 8 VVLTGAPTYVLAAG-----VWRLSAFLPARFYQALDRLCYVOSMVLFFPENTGYQ 62
DB 1 MVLVSTAFMMLIFWGLSAVLRIFS---IRSRKCVSFFSGSWALMPLEKINKTK 56
QY 63 ILIYGLPKKNENIYLANHSTVDWIVADILAIQNALGHVRYVLEKGLKMLPLYGCYF 122
DB 57 VIFSGDKVPCEDRVLLIANHRTVDWVYFMDLALRKQIGNIKYVLSLMKLPFGWAF 116
QY 123 AOHGGLYVRSKAFENKEMKNLQSYVDAGTPMYLVIFPEGTRYNPEQTKVLSAQAFAA 182
DB 117 HLEFEFIVERMVEDEANLRQIVSSFKDPDRLMLALFPEGTDYTEAKCO---RSKKFAA 173
```

```
QY 183 ORGLAVLKHYLTPRIKATHVAFDCMKNYLDIAYDVTVYE 222
DB 174 ENGLPILNNVLLPRTKGVSCLQELSCSDAVYDVITIGYK 213
```

```
RESULT 12
US-09-935-625-14202
; Sequence 14202, Application US/09935625
; GENERAL INFORMATION:
; APPLICANT: N. ALEXANDROV et al.
; TITLE OF INVENTION: POLYNUCLEOTIDES, POLYPEPTIDES, CELLS, AND METHODS THEREOF CAP
; FILE REFERENCE: 2750-1481P
; CURRENT APPLICATION NUMBER: US/09/935,625
; CURRENT FILING DATE: 2001-08-24
; NUMBER OF SEQ ID NOS: 33136
; SEQ ID NO 14202
; LENGTH: 351
; TYPE: PRP
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; NAME/KEY: peptide
; LOCATION: 1..351
; OTHER INFORMATION: Ceres Seq. ID no. 3441319
US-09-935-625-14202
```

```
Query Match
Best Local Similarity 23.9%; Score 288; DB 5; Length 351;
Matches 70; Conservative 47; Mismatches 91; Indels 12; Gaps 3;
```

```
QY 8 VVLTGAPTYVLAAG-----VWRLSAFLPARFYQALDRLCYVOSMVLFFPENTGYQ 62
DB 1 MVLVSTAFMMLIFWGLSAVLRIFS---IRSRKCVSFFSGSWALMPLEKINKTK 56
QY 63 ILIYGLPKKNENIYLANHSTVDWIVADILAIQNALGHVRYVLEKGLKMLPLYGCYF 122
DB 57 VIFSGDKVPCEDRVLLIANHRTVDWVYFMDLALRKQIGNIKYVLSLMKLPFGWAF 116
QY 123 AOHGGLYVRSKAFENKEMKNLQSYVDAGTPMYLVIFPEGTRYNPEQTKVLSAQAFAA 182
DB 117 HLEFEFIVERMVEDEANLRQIVSSFKDPDRLMLALFPEGTDYTEAKCO---RSKKFAA 173
QY 183 ORGLAVLKHYLTPRIKATHVAFDCMKNYLDIAYDVTVYE 222
DB 174 ENGLPILNNVLLPRTKGVSCLQELSCSDAVYDVITIGYK 213
```

```
RESULT 13
US-09-935-625-14255
; Sequence 14255, Application US/09935625
; GENERAL INFORMATION:
; APPLICANT: N. ALEXANDROV et al.
; TITLE OF INVENTION: POLYNUCLEOTIDES, POLYPEPTIDES, CELLS, AND METHODS THEREOF CAPA
; FILE REFERENCE: 2750-1481P
; CURRENT APPLICATION NUMBER: US/09/935,625
; CURRENT FILING DATE: 2001-08-24
; NUMBER OF SEQ ID NOS: 33136
; SEQ ID NO 14255
; LENGTH: 351
; TYPE: PRP
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; NAME/KEY: peptide
; LOCATION: 1..351
; OTHER INFORMATION: Ceres Seq. ID no. 3443249
US-09-935-625-14255
```

```
Query Match
Best Local Similarity 23.9%; Score 288; DB 5; Length 351;
Matches 70; Conservative 47; Mismatches 91; Indels 12; Gaps 3;
```

```

: SEQ ID NO 23026
: LENGTH: 351
: TYPE: PRT
: ORGANISM: Arabidopsis thaliana
: FEATURE:
: NAME/KEY: peptide
: LOCATION: 1..351
: OTHER INFORMATION: Ceres Seq. ID no. 3443249
US-09-935-625-23026

Query Match          23.9%; Score 288; DB 5; Length 351;
Best local similarity 31.8%; Pred. No. 2,3e-21;
Matches 70; Conservative 47; Mismatches 91; Indels 12; Gaps 3

QY      8  VLLGTAFTYYLWANG-----VWRLLSAEFLPARFYALDRLCYVYQSNYLFFFEENTGYQ 62
      :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
DB      1  MVLVSTAFMMLIFMGFLSAVVLRLFS----IYSRKCVSFFGSGMLALPFLFEKINKTK 56

QY      63  ILLIGDLPKKNENITYLANHOSYVMYIADIIATIQNALGHNRYVYLKEGLKPLPYGCF 122
      :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
DB      57  VIFSGDKVPCEDRYVLLINHRTVEYVMYFMDLALRKGGDIGNIKYVLLKSLMLPLFGMAF 116

QY      123  AQHGCIYVRSKAKFENEKEMRNKLOSIVYVAGFPMYLVIFPEGTRYVPEQIKVLSASOAFPA 182
      :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
DB      117  HLEETIPERKMEVDEANLRQIVSSFKDPRDALMLALFPEEGTIDYEAKQ---RSKKFPA 173

QY      183  QRGIAVLKHLVLPRIKATHVAFDCKNKNNIDAIYDVTVYE 222
      :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
DB      174  ENGLPLINNVLLPRTKGFVSCLOELSCSLDAVYDVITGYK 213

Search completed: August 28, 2002, 11:15:42
Job time: 556 sec

```

Search completed: August 28, 2002, 11:15:42  
Job time: 556 sec



---

**THIS PAGE BLANK (USPTO)**

nsen, N.F.; Hughes, B.; Huizar, L.

Nature 408, 816-820, 2000

A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maitl, R.; Marziani, Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.  
A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon, ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.  
A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.  
A:Reference number: A86141; MUID:21016719

A:Accession: B96780

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-393 <STO>

A:Cross-references: GB:AE005173; NID:96646762; PIDN:AAF21074.1; GSPDB:GN00141

C:Gene: F9E10.13

A:Map position: 1

#### Query Match

Best Local Similarity 32.3%; Score 312; DB 2; Length 393;  
Matches 76; Conservative 46; Mismatches 85; Indels 28; Gaps 4;

1 MRYLLPSVLTGTAPTVYLANGVRLLSAFLPARYQALDRLYCYQSMV----- 52

23 LRGLMILVFLSTAFMFL-----YFAP---IAALGLRLSVQOSKRVVSLIGLW 70

53 -----FFENTGYOILLYGDLPRKNENIYLANHOSTVDIVADILAIKONALGHVRY 107

71 LALPYLEFETVNGTIVFESGIIPEKRVLLIANHREVDIMYMLNLAIRKGCGLYKYV 130

108 LKELKMLPLCYGCFYAOHGITYKRSKFNKRNKLOSYVDAGTPMYLVIFPEGTTRYN 167

131 LKSLMLKLPFGWCFHLYLEFPERKREVDPEVLQMLSSFKDQDEPLMLALFPEGDTFT 190

168 PEQKVLASAOFAAORGALVAKHVLPRKATGVAFDCKMKNYDAITYDVTYVE 222

191 EEKCK---RSOKFAEYGLPALSNVLPKTRGFCVLEVLHNSIDAVYDITATK 242

Db 191 EEKCK---RSOKFAEYGLPALSNVLPKTRGFCVLEVLHNSIDAVYDITATK 242

RESULT 3

552645

probable 1-acyl-glycerol-3-phosphate acyltransferase - maize

C:Species: Zea mays (maize)

C>Date: 14-Jul-1995 #sequence\_revision 21-Jul-1995 #text\_change 19-May-2000

C:Accession: S52645

R:Brown, A.P.; Coleman, J.; Tommey, A.M.; Watson, M.D.; Slabas, A.R.

A:Title: Isolation and characterisation of a maize cDNA that complements a 1-acyl sn-gly-

her acyltransferases.

A:Reference number: S52645; MUID:95035993

A:Accession: S52645

A:Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-374 <BRO>

A:Cross-references: EMBL:Z29518; NID:9575959; PIDN:CAA82638.1; PID:9575960

C:Superfamily: Probable membrane protein YBR042c

#### Query Match

Best Local Similarity 31.7%; Score 269; DB 2; Length 374;  
Matches 71; Conservative 50; Mismatches 77; Indels 26; Gaps 6;

18 VLAWGVRLLSA-----FLPAR-----FYQALDRLYCYQSMVFFENTGYQV 62

9 VLPGLGLFLLSLVNAIQAVLFTIRPFSFYRIRINFLAELMLQIVAVVWMAAGVK 68

63 ILVYGD-----LPKKNENIYLANHOSTVDIVADILAIKONALGHVRYLKEGLKMLPLY 118

69 VOLHADEETYSKMGKHALIISNRSDIMLIGWILAQRSGCLSTLAVMKSSKFLPVI 128

119 G--CYFMOHGIIYKRSKFNKRNKLOSYVDAGTPMYLVIFPEGTTRYNPEGTXYLSA 176

129 GWSMFEAEY--LFLERSWADKDTKLKGLQRLKDFPRPFWLALFVEGTRETPAK--LLA 183

QY 177 SQAFPAORGIALVILKHLVLPRIKATVAFDCKMKNYDAITYDVTYV 220

Db 184 AOEFAASQGLPAPRNVLIPRTKGFVSASVIMRDEVPATYDTVI 227

#### RESULT 4

560478

probable 1-acyl-sn-glycerol-3-phosphate acyltransferase - Limnanthes douglasii

C:Species: Limnanthes douglasii

C>Date: 19-Mar-1997 #sequence\_revision 19-Mar-1997 #text\_change 17-Mar-2000

C:Accession: S60478

R:Brown, A.P.; Brough, C.L.; Kroon, J.T.M.; Slabas, A.R.

A:Title: Identification of a cDNA that encodes a 1-acyl-sn-glycerol-3-phosphate acylt

A:Reference number: S60477; MUID:96046746

A:Accession: S60478

A:Status: nucleic acid sequence not shown; translation not shown

A:Molecule type: mRNA

A:Residues: 1-377 <BRO>

A:Cross-references: EMBL:Z48730; NID:91067137; PIDN:CAA88620.1; PID:91067138

C:Note: the nucleotide sequence was submitted to the EMBL Data Library, March 1995

#### Query Match

Best Local Similarity 30.7%; Score 238; DB 2; Length 377;  
Matches 65; Conservative 49; Mismatches 74; Indels 24; Gaps 7;

26 LLSAFLPARY---QALDRLYCYQSMV-----LFFENTGYOILLYGDLPR--- 71

21 LVNVEIQAVFVVLVPRSPKDYRRIRNTLVALLMLLELVWIDMWAGVQVLYDTESEFL 80

72 -NKENIYLANHOSTVDIVADILAIKONALGHVRYLKEGLKMLPLYG--CYFAOHGI 128

81 MKRHALLICHRSDIDMLIGWYLAQRGCLSSIAVMKSSKFLPYIGSMFSEF--L 138

129 YKRSKFNKRNKLOSYVDAGTPMYLVIFPEGTTRYNPEQTKV--LSASQAFPAORGIA 187

139 FLERNMAKDEWTLKSGIQLRINDFPKPFMLALFVGTFRF---TKAKLLAAGEVAASGLP 194

QY 188 VLKHLVLPRIKATVAFDCKMKNYDAITYDVTYV 219

Db 195 VPRNVLLIPRTKGFVSASVIMRDEVPATYDTVI 226

#### RESULT 5

D96550

hypothetical protein FLIM15.12 [imported] - Arabidopsis thaliana

C:Species: Arabidopsis thaliana (mouse-ear cress)

C>Date: 02-Mar-2001 #sequence\_revision 02-Mar-2001 #text\_change 31-Mar-2001

C:Accession: D96550

R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alon

Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar,

anssen, N.F.; Hughes, B.; Hultzer, L.

A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim,

C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maitl, R.; Marzia

Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.

A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallo

ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.

A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.

A:Reference number: A86141; MUID:21016719

A:Accession: D96550

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-376 <STO>

A:Cross-references: GB:AE005173; NID:94836936; PIDN:AAD30638.1; GSPDB:GN00141

C:Gene: FLIM15.12

A:Map position: 1

```

Query Match          17.7%; Score 213.5; DB 2; Length 376;
Best Local Similarity 29.7%; Pred. No. 7.7e-12;
Matches 52; Conservative 46; Mismatches 66; Indels 11; Gaps 4;

OY 51 VLFFPENTGVOILLYGDLPK----NKENIITLANHOSVDIMVADILAIKRONALGHVY 106
   ::::: ::::: ::::: ::::: ::::: ::::: ::::: ::::: ::::: :::::
Db 57 LITLFDWMACIKINLYVAETLELICKKEHALVLSNHRSDIDWLIGVMAQVCGSSIA 116

OY 107 VLKEELKMLPLUG--CYFAQHGCIYVRSKAKNEKEMRNKLOSVDACGPMVLVPEECT 164
   ::::: ::::: ::::: ::::: ::::: ::::: ::::: ::::: ::::: :::::
Db 117 IMKKRAKTLPIITIGMSMWPSDY--IFLEKSWAKDENTLLKGFRLDEFPMTFMLALFEVET 174

OY 165 RYNPOTKVLASQAFAAORGIAVLKHLVTPRIKATHVAFDCMKNYLDAIVDTV 219
   ::::: ::::: ::::: ::::: ::::: ::::: ::::: ::::: ::::: :::::
Db 175 RTGOKK---LEAAQEVASIRSLPSPRNVLIPRTKGFVSAVSRVIRSVPAIVDTCTL 226

RESULT 6
T31913
hypothetical protein T05H4.1 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 29-Oct-1999
C:Accession: T31913
R:Blanchard, M.
submitted to the EMBL Data Library, July 1997
A:Description: The sequence of C. elegans cosmid T05H4.
A:Reference number: 221097
A:Accession: T31913
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-344 <BLA>
A:Cross-references: EMBL:AF016452; PIDN:AA66008.1; GSPDB:GN00023; CESP:T05H4.1
A:Experimental source: strain Bristol N2; clone T05H4
C:Genetics:
A:Gene: CESP:T05H4.1
A:Map position: 5
A:Introns: 55/3; 212/1; 243/3; 298/3

Query Match          16.6%; Score 200; DB 2; Length 344;
Best Local Similarity 28.7%; Pred. No. 1.2e-10;
Matches 66; Conservative 35; Mismatches 93; Indels 36; Gaps 10;

OY 10 LLGTG----PTVVLAMGWRLLSAFLPAFYQALDDRLCYVQSMWLFPEYNTGVQILL 65
   ||||| ::::: ::::: ::::: ::::: ::::: ::::: ::::: ::::: :::::
Db 18 LLGTFLFLPLPLPLAM-----FAP-KLMRTCADRLVGFMLTPPCSLIEVVFQYNEFY 68

OY 66 YGDLPRKNENIITLANHOSVDIMVADILAIKRONALGHV-----RYVLEKGIKMLPL 117
   ||| ::::: ::::: ::::: ::::: ::::: ::::: ::::: ::::: :::::
Db 69 TGDILRDEPAILIMNHRRLDMLFS-----WNALYKMDPWLITTEKISTAKPLKKIP- 121

OY 118 YGCTFAQHGCIYVRSKAF-NEKEMRNKLOSVDACGPMY-LVTFPEGTRVNEPOTKVL 175
   ::::: ::::: ::::: ::::: ::::: ::::: ::::: ::::: :::::
Db 122 -GAGNAMSGSYIFLDRNFENDKPVLERIVKYSGSEKKYQILLFAEGDKGERATRL-- 178

OY 176 ASQAFAAORGIAVLKHLVTPRIKATHVAFDCMK--NYLDAIVDVPVVEG 223
   ||| ::::: ::::: ::::: ::::: ::::: ::::: ::::: ::::: :::::
Db 179 -SDAFADKNGLPREYEVLPHTTGKFKFLMELMKKENIYVDTLTIANS 227

RESULT 7
S54641
probable membrane protein YDR018c - yeast (Saccharomyces cerevisiae)
N:Alternate names: hypothetical protein D3246; hypothetical protein PZF396; hypothetical
C:Species: Saccharomyces cerevisiae
C:Date: 08-Jul-1995 #sequence_revision 01-Sep-1995 #text_change 20-Jun-2000
C:Accession: S54641; S63425; S67831; S72116
R:Deedman, K.; Brown, D.; Hamlyn, N.; Bowman, S.
submitted to the EMBL Data Library, May 1995
A:Reference number: S54638
A:Accession: S54641
A:Molecule type: DNA
A:Residues: 1-396 <DED>

A:Cross-references: EMBL:Z49770; NID:g840867; PIDN:CAA89843.1; PID:g840871
A:Experimental source: strain AB972
R:Eide, L.G.; Sander, C.; Prydz, H.
submitted to the EMBL Data Library, February 1996
A:Description: Sequencing and analysis of a 35.4 kb region on the left arm of chromos
A:Reference number: S63416
A:Accession: S63425
A:Molecule type: DNA
A:Residues: 1-396 <EID>
A:Cross-references: EMBL:X95966; NID:g1216215; PIDN:CAA65210.1; PID:g1216225
R:Prydz, H.; Eide, L.G.
submitted to the Protein Sequence Database, July 1996
A:Reference number: S67822
A:Accession: S67831
A:Molecule type: DNA
A:Residues: 1-396 <PRY>
A:Cross-references: EMBL:Z74314; NID:g1431443; PIDN:CAA98838.1; PID:g1431444; MIPS:YD
A:Experimental source: strain S288C
R:Eide, L.G.; Sander, C.; Prydz, H.
Yeast 12. 1085-1090, 1996
A>Title: Sequencing and analysis of a 35.4 kb region on the left arm of chromosome IV
A:Reference number: S72107; MUID:97051598
A:Accession: S72116
A>Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-396 <EIM>
A:Cross-references: EMBL:X95966; NID:g1216215; PIDN:CAA65210.1; PID:g1216225
A>Note: the nucleotide sequence was submitted to the EMBL Data Library, February 1996
C:Genetics:
A:Map position: 4R
A>Note: YDR018c
C:Superfamily: probable membrane protein YBR042c
C:Keywords: transmembrane protein
E:27-43/Domain: transmembrane #status predicted <TM1>
E:69-85/Domain: transmembrane #status predicted <TM2>
E:376-392/Domain: transmembrane #status predicted <TM3>

Query Match          15.5%; Score 186.5; DB 2; Length 396;
Best Local Similarity 31.0%; Pred. No. 2.5e-09;
Matches 53; Conservative 26; Mismatches 67; Indels 25; Gaps 4;

OY 73 KENIITLANHOSVDIMVADILAIKRONALGHVYVLEKGIKMLPLCYFAQHGCIYVNR 132
   |:::| ::::: ::::: ::::: ::::: ::::: ::::: ::::: :::::
Db 108 KDRATITLANHOKYADPIYIMLSTFVSNLGGNYIILKRLQIIFPLGPEGMRKFKTFELSR 167

OY 133 SAKFNEKEMRNKL-----QSYVDACGPMY-LVTFPEGTRVNEPOTK 172
   ::::: ::::: ::::: ::::: ::::: ::::: ::::: :::::
Db 168 NMQKDEKALTNLSLVSMDLNARCKGPLTNYKSCYSKTNESIAVNLIMPEGTNLS---LK 224

OY 173 VLSAQAFAAORGCL-AVLKHLVTPRIKATHVAFDCMKNYLDAIVDTVYV 221
   ::::: ::::: ::::: ::::: ::::: ::::: ::::: :::::
Db 225 TREKSEAFQORALHDVQRLHLPLPSKGLKRAVEKLAPSLDAIVDTVIGY 275

RESULT 8
T07936
probable glycerol-3-phosphate O-acetyltransferase (EC 2.3.1.15) - rape
N:Alternate names: 1-acyl-sn-glycerol-3-phosphate acyltransferase
C:Species: Brassica napus (rape)
C:Date: 21-May-1999 #sequence_revision 21-May-1999 #text_change 20-Jun-2000
C:Accession: T07936
R:Brough, C.L.
submitted to the EMBL Data Library, June 1995
A:Reference number: Z16230
A:Accession: T07936
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-311 <BRO>
A:Cross-references: EMBL:Z49860; NID:g1149594; PIDN:CAA90019.1
A:Experimental source: cv. jet neuf; embryo
C:Superfamily: probable membrane protein YBR042c
C:Keywords: acyltransferase; coenzyme A

```





A:Experimental source: clone F55A11  
C:Gene(s):  
A:Gene: CESP:F55A11.5  
A:Map position: 5  
A:Introns: 19/3; 52/2; 176/1; 222/3; 260/2; 327/1; 342/1

Query Match 11.1%; Score 134; DB 2; Length 439;  
Best Local Similarity 23.6%; Pred. No. 0.00019;  
Matches 45; Conservative 40; Mismatches 74; Indels 32; Gaps 4;  
QY 60 GVOILLYGDLPKKNKENTITLANQSTVDWIV--ADILAIRONALGHVRYVLEKGLKWLPL 117  
Db 27 GVRIRVSGDEIERGSPAMIVMNHRTLDWMYMCALYQINPWLITSNKSISLKAQLKRLPG 86  
QY 118 YGCFYAOHGSIYVRSKAFNEKEKRNKLOSVDAGTPMVLVIFPEGTRVNPQOTKVLAS 177  
Db 87 AGFGMAAAGVFLERNAEVDKRSFDDAIDYFKNIDKKYQILLEPEGTDKSEWTT---LKS 143  
QY 178 QAFPAQOGLAVLKHVLTFR-----IKATHVAF-----DCMKNY 210  
Db 144 REFPAKKNGLRHLDYVLYPPTTGLHLNKMREPSSVIRFKNVHFYRFIKIQKIIILEEY 203  
QY 211 LDAIYDVTVYV 221  
Db 204 VEXIYDITIAV 214

Search completed: August 28, 2002, 11:08:59  
Job time: 253 sec



GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 28, 2002, 11:17:23 ; Search time 21.99 seconds

(without alignments)  
401.458 Million cell updates/sec

Title: US-09-853-526-70

Perfect score: 1203

Sequence: 1 MRYLLPSVVLGTAPTYVLA.....NYLDAIVYTVVEGKDDGC 228

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 105224 seqs, 38719550 residues

Total number of hits satisfying chosen parameters: 105224

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Swissprot\_40:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1203	100.0	353	1 P1CE_HUMAN	09nq2 homo sapien
2	1040	86.5	354	1 P1CE_MOUSE	09d1e8 homo musculi
3	197.5	16.4	378	1 P1CD_RAT	092s1 rattus norv
4	196	16.3	378	1 P1CD_HUMAN	09n25 homo sapien
5	187.5	15.6	376	1 P1CC_HUMAN	09n27 homo sapien
6	186.5	15.5	396	1 YD18_YEAST	012185 saccharomyc
7	161	13.4	397	1 YBP2_YEAST	P38226 saccharomyc
8	125.5	10.4	370	1 Y205_HUMAN	092604 homo sapien
9	120.5	10.0	303	1 P1SC_YEAST	P33333 saccharomyc
10	120.5	10.0	310	1 Y1HG_ECOLI	P32129 escherichia
11	119	9.9	281	1 P1SC_LIMO	042870 limnanthes
12	117.5	9.8	391	1 Y173_CAEEL	011087 caenorhabdi
13	116	9.6	281	1 P1SC_LIMAL	042868 limnanthes
14	114	9.5	266	1 P1SC_MYCPN	P75479 mycoplasma
15	111	9.2	377	1 VENV_FOWPV	P36316 fowlpox vir
16	109	9.1	240	1 P1SC_HAELN	P44848 haemophilus
17	104	8.6	240	1 P1SC_LIMO	025903 helicobacte
18	101	8.4	308	1 P1SC_COCON	042670 cocco nucif
19	99.5	8.3	237	1 P1SC_HELPI	0921n8 helicobacte
20	99	8.2	245	1 P1SC_SALTY	P261n8 salmonella
21	97.5	8.1	262	1 P1C1_CAEEL	093841 caenorhabdi
22	97.5	8.1	488	1 EXON_HSV6U	P24447 human herpe
23	97	8.1	245	1 P1SC_ECOLI	022267 escherichia
24	94.5	7.9	282	1 P1C2_CAEEL	022267 caenorhabdi
25	94.5	7.9	488	1 EXON_HSV6Z	P52448 human herpe
26	93	7.7	250	1 P1SC_BORBU	059188 borrelia bu
27	90	7.5	484	1 Y142_NPVOB	016370 oryza pseu
28	87.5	7.3	446	1 M164_MOUSE	061542 mus musculi
29	83.5	6.9	268	1 P1SC_MYCGE	024902 mycoplasma
30	82	6.8	503	1 CPN1_PAPHA	025527 papilo hamad
31	80	6.7	529	1 YDUL_CAEEL	P30638 caenorhabdi
32	79.5	6.6	2156	1 RRP1_PUDMH	P27176 pumila vir
33	79	6.6	281	1 Y1XC_BACSU	P06569 bacillus su

34	78.5	6.5	204	1 NDCM_RECAM	021271 reclinomona
35	78	6.5	445	1 M164_HUMAN	014849 homo sapien
36	78	6.5	503	1 CPN1_HUMAN	P15538 homo sapien
37	77.5	6.4	371	1 Y1B0_CLOAB	004354 clostridium
38	77.5	6.4	383	1 FDEE_ARATH	P46313 arabidopsis
39	77	6.4	999	1 DSG3_HUMAN	P32926 homo sapien
40	76.5	6.4	283	1 P1CA_HUMAN	099943 homo sapien
41	76.5	6.4	531	1 UD16_MOUSE	064435 mus musculi
42	76	6.3	307	1 SYGA_XY1FA	09pc25 xyella fas
43	76	6.3	384	1 FDEE_BRAUO	039287 brassica ju
44	76	6.3	892	1 RAI6_SCHPO	P36617 schizosacch
45	75.5	6.3	326	1 305B_RAT	P31210 rattus norv

## ALIGNMENTS

RESULT	ID	PLCE_HUMAN	STANDARD	PRT	353 AA.
AC	09NIO2: 09B0G4:				
DT	01-MAR-2002 (Rel. 41, Created)				
DT	01-MAR-2002 (Rel. 41, Last sequence update)				
DT	01-MAR-2002 (Rel. 41, Last annotation update)				
DE	1-acyl-sn-glycerol-3-phosphate acyltransferase epsilon (EC 2.3.1.51)				
DE	(1-AGP acyltransferase 5) (1-AGPAT 5) (Lysophosphatidic acid				
DE	acyltransferase-epsilon) (LPAAT-epsilon) (1-acylglycerol-3-phosphate				
DE	0-acyltransferase 5).				
GN	AGPAT5.				
OS	Homo sapiens (Human).				
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;				
OX	Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.				
OX	NCBI_TaxID=9606;				
RP	SEQUENCE FROM N.A.				
RA	Leung D.W.;				
RT	"Cloning and expression of LPAAT-epsilon.";				
RL	Submitted (MAY-2001) to the EMBL/Genbank/DDJ databases.				
RN					
RP	SEQUENCE FROM N.A.				
RA	Cohen D., Chumakov I., Blumenfeld M., Bougueleret L.;				
RL	Patent number WO932644, 01-JUL-1999.				
RN	[3]				
RP	SEQUENCE FROM N.A.				
RA	TISSUE=amygdala;				
RX	MEDLINE=21154917; PubMed=11230166;				
RA	Wiemann S., Well B., Wellenreuther R., Gassenhuber J., Glassl S.,				
RA	Ansoerg W., Boecker M., Bloeker H., Bauersachs S., Blum H.,				
RA	Lauber J., Duesterhoeft A., Beyer A., Koehrer K., Strack N.,				
RA	Mewes H.-W., Oltmawelder B., Obermayer B., Tampe J., Heubner D.,				
RA	Wambolt R., Korn B., Klein M., Poustka A.;				
RT	"Towards a catalog of human genes and proteins: sequencing and				
RT	analysis of 500 novel complete protein coding human cDNAs.";				
RT	Genome Res. 11:422-435(2001).				
RN	[4]				
RP	SEQUENCE FROM N.A.				
RA	TISSUE=Placenta;				
RC					
RA	Isogai T., Ota T., Hayashi K., Sugiyama T., Otsuki T., Suzuki Y.,				
RA	Nishikawa T., Nagai K., Sugano S., Shiratori A., Sudo H.,				
RA	Wagatsuna M., Hosokiri T., Kaku Y., Kodaira H., Kondo H., Sugawara M.,				
RA	Takahashi M., Chiba Y., Ishida S., Murakawa K., Ono Y., Takiguchi S.,				
RA	Watanabe S., Kimura K., Murakami K., Ishii S., Kawai Y., Saito K.,				
RA	Yamamoto J., Wakematsu A., Nakamura Y., Nagahara K., Masuno Y.,				
RA	Ninomiya K., Iwayanagi T.;				
RT	"NEO human cDNA sequencing project.";				
RT	Submitted (FEB-2000) to the EMBL/Genbank/DDJ databases.				
CC	-1- FUNCTION: CONVERTS LYSOPHOSPHATIDIC ACID (LPA) INTO PHOSPHATIDIC				
CC	ACID BY INCORPORATING ACYL MOIETY AT THE 2 POSITION (BY				
CC	SIMILARITY).				
CC	-1- CATALYTIC ACTIVITY: Acyl-CoA + 1-acyl-sn-glycerol 3-phosphate -				
CC	CoA + 1,2-diacyl-sn-glycerol 3-phosphate.				
CC	-1- PATHWAY: SECOND STEP IN DE NOVO PHOSPHOLIPID BIOSYNTHESIS.				
CC	-1- SUBCELLULAR LOCATION: Integral membrane protein (Potential).				

CC -1- SIMILARITY: BELONGS TO THE 1-ACYL-SN-GLYCEROL-3-PHOSPHATE  
 CC ACYLTRANSFERASE FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL: AF375789; AAK54809.1; ALT\_INIT.  
 DR EMBL: ALJ36587; CAB66522.1; ALT\_INIT.  
 DR EMBL: AK002072; BAA92069.1; -  
 DR InterPro: IPR002123; Acyltransferase.  
 DR Pfam: PF01553; Acyltransferase; 1.  
 KM Phospholipid biosynthesis; Transferase; Acyltransferase;  
 KM Transmembrane.  
 FT TRANSMEM 7 29 POTENTIAL.  
 FT TRANSMEM 44 66 POTENTIAL.  
 FT TRANSMEM 334 351 POTENTIAL.  
 FT CONFLICT 145 145 L -> V (IN REF. 2).  
 SO SEQUENCE 353 AA; 40813 MW; A05B1FA246CE1B64 CRC64;

Query Match 100.0%; Score 1203; DB 1; Length 353;  
 Best Local Similarity 100.0%; Pred. No. 3.3e-104;  
 Matches 228; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MRYLLPSVLLGTPAPTVLAWGWRLSAPFLPAFYQALDDRLCYQGSVLPFFENYTG 60  
 DB 1 MRYLLPSVLLGTPAPTVLAWGWRLSAPFLPAFYQALDDRLCYQGSVLPFFENYTG 60  
 QY 61 VOILLVGDLPKKNENIYLLANHOSTVDWIVADIATRONALGHVRYLKEGLKMLPLPGC 120  
 DB 61 VOILLVGDLPKKNENIYLLANHOSTVDWIVADIATRONALGHVRYLKEGLKMLPLPGC 120  
 QY 121 YFAOHGCIYVRSKAKFNEKEMRNKLOSVDAGTPMYLVIFPEGTRYNPEQTKVLSAQAF 180  
 DB 121 YFAOHGCIYVRSKAKFNEKEMRNKLOSVDAGTPMYLVIFPEGTRYNPEQTKVLSAQAF 180  
 QY 181 AAOAGLAVLKHVLTLPRIKATHVAFDCKKNYLDIYDVTVYEEGKDDG 228  
 DB 181 AAOAGLAVLKHVLTLPRIKATHVAFDCKKNYLDIYDVTVYEEGKDDG 228

RESULT 2  
 PLCD\_MOUSE ID PLCD\_MOUSE STANDARD; PRT; 354 AA.

AC 09D1E8;  
 DT 01-MAR-2002 (Rel. 41, Created)  
 DT 01-MAR-2002 (Rel. 41, Last sequence update)  
 DT 01-MAR-2002 (Rel. 41, Last annotation update)  
 DE 1-acyl-sn-glycerol-3-phosphate acyltransferase epsilon (EC 2.3.1.51)  
 DE (1-AGP acyltransferase 5) (1-AGPAT 5) (Lysophosphatidic acid  
 DE acyltransferase-epsilon) (LPAAT-epsilon) (1-acylglycerol-3-phosphate  
 DE O-acyltransferase 5).  
 DE GN AGPAT5.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 ON NCBI\_TaxID=10090;  
 RN (1)  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=C57BL/6J; TISSUE=Embryo;  
 RX MEDLINE=21085660; PubMed=11217851;  
 RA Kawal J., Shingawa A., Shidata K., Yoshino M., Itoh M., Ishii Y.,  
 RA Aizawa K., Izawa A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,  
 RA Saito T., Okazaki Y., Nishi K., Kiyosawa H., Kondo S., Yamana K.,  
 RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,  
 RA Katsuda K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,  
 RA Fleischmann W., Gaasterland T., Glass C., King B., Kochiwa H.,  
 RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,

RA Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,  
 RA Sakai K., Okido T., Furuno M., Aono H., Balderelli R., Barsh G.,  
 RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,  
 RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,  
 RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamuya M., Lee N.H.,  
 RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,  
 RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,  
 RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,  
 RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whitaker C., Wilmink L.,  
 RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohsaki S.,  
 RA Hayashizaki Y.;  
 RT "Functional annotation of a full-length mouse cDNA collection.";  
 RL Nature 409:685-690(2001).  
 CC -1- FUNCTION: CONVERTS LYSOPHOSPHATIDIC ACID (LPA) INTO PHOSPHATIDIC  
 CC ACID BY INCORPORATING ACYL MOIETY AT THE 2 POSITION (BY  
 CC SIMILARITY).  
 CC -1- CATALYTIC ACTIVITY: Acyl-CoA + 1-acyl-sn-glycerol 3-phosphate -  
 CC CoA + 1,2-diacyl-sn-glycerol 3-phosphate.  
 CC -1- PATHWAY: SECOND STEP IN DE NOVO PHOSPHOLIPID BIOSYNTHESIS.  
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein (Potential).  
 CC -1- SIMILARITY: BELONGS TO THE 1-ACYL-SN-GLYCEROL-3-PHOSPHATE  
 CC ACYLTRANSFERASE FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL: AK003649; BAB22915.1; -  
 DR MGD: MGI:1915880; Agpat5.  
 DR InterPro: IPR002123; Acyltransferase.  
 DR Pfam: PF01553; Acyltransferase; 1.  
 KM Phospholipid biosynthesis; Transferase; Acyltransferase;  
 KM Transmembrane.  
 FT TRANSMEM 7 29 POTENTIAL.  
 FT TRANSMEM 44 66 POTENTIAL.  
 FT TRANSMEM 335 352 POTENTIAL.  
 SO SEQUENCE 354 AA; 40943 MW; 447EEF924B91E800 CRC64;

Query Match 86.5%; Score 1040; DB 1; Length 354;  
 Best Local Similarity 85.9%; Pred. No. 4.3e-89;  
 Matches 195; Conservative 19; Mismatches 13; Indels 0; Gaps 0;

QY 1 MRYLLPSVLLGTPAPTVLAWGWRLSAPFLPAFYQALDDRLCYQGSVLPFFENYTG 60  
 DB 1 MRYLLPSVLLGTPAPTVLAWGWRLSAPFLPAFYQALDDRLCYQGSVLPFFENYTG 60  
 QY 61 VOILLVGDLPKKNENIYLLANHOSTVDWIVADIATRONALGHVRYLKEGLKMLPLPGC 120  
 DB 61 VOILLVGDLPKKNENIYLLANHOSTVDWIVADIATRONALGHVRYLKEGLKMLPLPGC 120  
 QY 121 YFAOHGCIYVRSKAKFNEKEMRNKLOSVDAGTPMYLVIFPEGTRYNPEQTKVLSAQAF 180  
 DB 121 YFAOHGCIYVRSKAKFNEKEMRNKLOSVDAGTPMYLVIFPEGTRYNPEQTKVLSAQAF 180  
 QY 181 AAOAGLAVLKHVLTLPRIKATHVAFDCKKNYLDIYDVTVYEEGKDDG 227  
 DB 181 AAOAGLAVLKHVLTLPRIKATHVAFDCKKNYLDIYDVTVYEEGKDDG 227

RESULT 3  
 PLCD\_RAT ID PLCD\_RAT STANDARD; PRT; 378 AA.

AC 0924S1;  
 DT 01-MAR-2002 (Rel. 41, Created)  
 DT 01-MAR-2002 (Rel. 41, Last sequence update)  
 DT 01-MAR-2002 (Rel. 41, Last annotation update)  
 DE 1-acyl-sn-glycerol-3-phosphate acyltransferase delta (EC 2.3.1.51) (1-  
 DE AGP acyltransferase 4) (1-AGPAT 4) (Lysophosphatidic acid

DE acyltransferase-delta) (LPAAT-delta) (1-acylglycerol-3-phosphate O-  
acetyltransferase 4).  
GN AGPAT4.  
OS Rattus norvegicus (Rat).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
OX NCBI\_TaxID=10116;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Li W., Suzuki T.:  
RT "Rattus norvegicus mRNA for lysophosphatidic acid  
acetyltransferase-delta, complete cds."  
RL Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.  
CC -1- FUNCTION: CONVERTS LYSPHOSPHATIDIC ACID (LPA) INTO PHOSPHATIDIC  
CC ACID BY INCORPORATING ACYL MOIETY AT THE 2 POSITION (BY  
CC SIMILARITY).  
CC -1- CATALYTIC ACTIVITY: Acyl-CoA + 1-acyl-sn-glycerol 3-phosphate =  
CC CoA + 1,2-diacyl-sn-glycerol 3-phosphate.  
CC -1- PATHWAY: SECOND STEP IN DE NOVO PHOSPHOLIPID BIOSYNTHESIS.  
CC -1- SUBCELLULAR LOCATION: Integral membrane protein (potential).  
CC -1- SIMILARITY: BELONGS TO THE 1-ACYL-SN-GLYCEROL-3-PHOSPHATE  
CC ACYLTRANSFERASE FAMILY.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL: AB067572; BAB62290.1; -  
KM Phospholipid biosynthesis; Transferase; Acyltransferase;  
KW Transmembrane  
FT TRANSMEM 11 31 POTENTIAL.  
FT TRANSMEM 125 145 POTENTIAL.  
FT TRANSMEM 307 327 POTENTIAL.  
FT TRANSMEM 338 358 POTENTIAL.  
SQ SEQUENCE 378 AA; 43794 MW; 389AA01B7327AE2B CRC64;

Query Match 16.4%; Score 197.5; DB 1; Length 378;  
Best Local Similarity 27.4%; Pred. No. 6.2e-11;  
Matches 55; Conservative 46; Mismatches 85; Indels 15; Gaps 7;  
QY 34 REFQALDDL-KYVYSMWLEFFENTGYOILLYGDLPR-----KNEITYLANHSTYD 87  
DB 44 QLEFKINARCYCV-SSQVLMLEWMSGTECTITYD-PRASPHYGKEMALIVLNHKEID 101  
QY 88 WIYADILAIRONALGHVRYVLEKGLKMLPLYG--CYFAOHGGIYVRSKAFNEKEMRNKL 145  
DB 102 FLGNSLAERLGLSKYLAKELAYVPLIGMMYFVEH--IFCRKWEQDQOTYAKSL 159  
QY 146 QSYVDAGTWMYLVPEEGTRYNEQOTYVLSASQAFARGLAVLKHYLPRIKATHVAFD 205  
DB 160 LHLRDYPERKYLFIHCEGRTFTEKKHDI--SMQVAAQAGLPRLKHHLLPRTGFAITYK 216  
QY 206 CMKNYDAITYDYTVYVEGKDD 226  
DB 217 CLRDVPAVAYDCTLNFRRNEN 237

RESULT 4  
ID PLCC\_HUMAN STANDARD; PRT; 378 AA.  
AC Q9NR25;  
DT 16-OCT-2001 (Rel. 40, Created)  
DT 16-OCT-2001 (Rel. 40, Last sequence update)  
DT 16-OCT-2001 (Rel. 40, Last annotation update)  
DE 1-acyl-sn-glycerol-3-phosphate acyltransferase delta (EC 2.3.1.51) (1-  
AGP acyltransferase 4) (1-AGPAT 4) (lysophosphatidic acid  
acyltransferase-delta) (LPAAT-delta) (1-acylglycerol-3-phosphate O-  
acetyltransferase 4).

GN AGPAT4.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Leung D.W.;  
RT "Structure and functions of lysophosphatidic acid acyltransferases";  
RL Submitted (JUN-1999) to the EMBL/GenBank/DBJ databases.  
CC -1- FUNCTION: CONVERTS LYSPHOSPHATIDIC ACID (LPA) INTO PHOSPHATIDIC  
CC ACID BY INCORPORATING ACYL MOIETY AT THE 2 POSITION (BY  
CC SIMILARITY).  
CC -1- CATALYTIC ACTIVITY: Acyl-CoA + 1-acyl-sn-glycerol 3-phosphate =  
CC CoA + 1,2-diacyl-sn-glycerol 3-phosphate.  
CC -1- PATHWAY: SECOND STEP IN DE NOVO PHOSPHOLIPID BIOSYNTHESIS.  
CC -1- SUBCELLULAR LOCATION: Integral membrane protein (potential).  
CC -1- SIMILARITY: BELONGS TO THE 1-ACYL-SN-GLYCEROL-3-PHOSPHATE  
CC ACYLTRANSFERASE FAMILY.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL: AF156776; AAF80338.1; -  
DR InterPro: IPR002123; Acyltransferase.  
KW Pfam: PF01553; Acyltransferase; 1.  
KM Phospholipid biosynthesis; Transferase; Acyltransferase;  
KW Transmembrane  
FT TRANSMEM 11 31 POTENTIAL.  
FT TRANSMEM 125 145 POTENTIAL.  
FT TRANSMEM 307 327 POTENTIAL.  
FT TRANSMEM 338 358 POTENTIAL.  
SQ SEQUENCE 378 AA; 44021 MW; 3EF013D196F8C0C5 CRC64;

Query Match 16.3%; Score 196; DB 1; Length 378;  
Best Local Similarity 24.8%; Pred. No. 8.6e-11;  
Matches 57; Conservative 55; Mismatches 96; Indels 22; Gaps 8;  
QY 5 LPSVLLGTAPRY-VLAWGVRLLSFLPRFQALDDL-KYVYSMWLEFFENTGYG 62  
DB 22 IASGLIINTYQLTLLPMPINKOL-----FRKINCRSLSTCI-SSQVLMLEWMSGTE 72  
QY 63 ILLYGD---LPRKNEITYLANHSTYDVIYADILAIRONALGHVRYVLEKGLKMLPLY 118  
DB 73 CTIFTPRALYKKSKENAIYVNLHKEIDPLCGMSLSERGLGSKYLAKRELAVPII 132  
QY 119 G--CYFAOHGGIYVRSKAFNEKEMRNKLQSYVDAGTWMYLVPEEGTRYNEQOTYVLSA 176  
DB 133 GMMWYFTEM--VPCSRKWEQDRKTVAATSLQHLRDYDEKYEFLIHCEGRTFTEKKHEI--- 187  
QY 177 SQAFARGLAVLKHYLPRIKATHVAFDQMKRYVLAITYDYTVYVEGKDD 226  
DB 188 SMQVAAQAGLPRLKHHLLPRTGFAITYVSLRNVAAYDCTLNFRRNEN 237

RESULT 5  
ID PLCC\_HUMAN STANDARD; PRT; 376 AA.  
AC Q9NR27; Q9NR26;  
DT 16-OCT-2001 (Rel. 40, Created)  
DT 16-OCT-2001 (Rel. 40, Last sequence update)  
DT 01-MAR-2002 (Rel. 41, Last annotation update)  
DE 1-acyl-sn-glycerol-3-phosphate acyltransferase gamma (EC 2.3.1.51) (1-  
AGP acyltransferase 3) (1-AGPAT 3) (lysophosphatidic acid  
acyltransferase-gamma) (LPAAT-gamma) (1-acylglycerol-3-phosphate O-  
acetyltransferase 3).  
GN AGPAT3.



```

Db      225  TREKSEARCCRAHLDHVLQHLRLHLLPHSGLKFAVEAKLAPSLDAIDVYITGY 275
RESULT 7
ID      YBP2_YEAST
AC      P38226;
DT      01-OCT-1994 (Rel. 30, Created)
DT      01-OCT-1994 (Rel. 30, Last sequence update)
DT      16-OCT-2001 (Rel. 40, Last annotation update)
DE      Hypothetical 45.5 kDa protein in FAY1-TCM62 intergenic region.
GN      YBR042C OR YBR0412.
OS      Saccharomyces cerevisiae (Baker's yeast).
OC      Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC      Saccharomycetales; Saccharomycetaceae; Saccharomyces.
OX      NCBI_Taxid:4932;
RN      [1]
RP      SEQUENCE FROM N.A.
RC      STRAIN:S288C;
RA      Andre B., Czjzekluch C., Hein C., Jauniaux J.-C., Urrastarazu A.,
RA      Vissers S.;
RL      Submitted (AUG-1994) to the EMBL/GenBank/DBJ databases.
CC      -1- SUBCELLULAR LOCATION: Integral membrane protein (Potential).
CC      -1- SIMILARITY: BELONGS TO THE 1-ACYL-SN-GLYCEROL-3-PHOSPHATE
CC      ACYLTRANSFERASE FAMILY.
CC      -----
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC      or send an email to license@isb-sib.ch).

```

[illegible]

DT	Q92604; 15-JUL-1999 (Rel. 38, Created)
DT	15-JUL-1999 (Rel. 38, Last sequence update)
DT	16-OCT-2001 (Rel. 40, Last annotation update)
DE	Hypothetical protein KIAA0205.
GN	KIAA0205.
OS	Homo sapiens (Human).
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX	NCBI_TaxID:9606;
RN	[1]
RP	SEQUENCE FROM N.A.
RC	TISSUE=Bone marrow;
RX	MEDLINE=97191544; PubMed=9039502;
RA	Nagase T., Seki N., Ishikawa K.-I., Ohira M., Kawarabayasi Y.,
RA	Ohara O., Tanaka A., Kotani H., Miyajima N., Nomura N.;
RT	"Prediction of the coding sequences of unidentified human genes. VI.
RT	The coding sequences of 80 new genes (KIAA0201-KIAA0280) deduced by
RT	analysis of cDNA clones from cell line KG-1 and brain.;"
RL	DNA Res. 3:321-329(1996).
CC	-1- SUBCELLULAR LOCATION: Integral membrane protein (Potential).
CC	-1- SIMILARITY: BELONGS TO THE 1-ACYL-SN-GLYCEROL-3-PHOSPHATE
CC	ACYLTRANSFERASE FAMILY.
CC	-----
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC	the European Bioinformatics Institute. There are no restrictions on its
CC	use by non-profit institutions as long as its content is in no way
CC	modified and this statement is not removed. Usage by and for commercial
CC	entities requires a license agreement (See <a href="http://www.isb-sib.ch/announce/or">http://www.isb-sib.ch/announce/</a>
CC	or send an email to <a href="mailto:license@isb-sib.ch">license@isb-sib.ch</a> ).
CC	-----
DR	EMBL; D86960; BAA13196.1; -
DR	InterPro; IPR002123; Acyltransferase.
KW	Hypothetical protein; Phospholipid biosynthesis; Transferase;
KW	Acyltransferase; Transmembrane.
FT	TRANSMEM 22 42
FT	TRANSMEM 342 362 POTENTIAL.
SO	SEQUENCE 370 AA; 43089 MW; 93496544EA651541 CRC64;

	Query Match	Similarity	10.4%	Score 125.5	DB: 1;	Length 370;	
	Best Local Similarity		22.58%;	Pred. No. 0.00029;			
	Matches	56;	Conservative	51;	Mismatches	105;	Indels 37; Gaps 11;
Oy	1	KRLLPSPVLLGTAPTYVLANGVMHLLSAPLPARYQALDRLYCYVOSMYLFEENYTG	60				
		: :   :   : : :   :   :   :   :   :   :					
Dd	20	MRFAFMVNNNLVAIPSYICYIILLDPLAVLDSSKRWY-IEGIMTKMLLGWA-SMGWAG	77				
Oy	61	VOILLYGDLPR--NKENITYLANHOSTVDWIADILAIRQN--ALGHVRVYLEGKWL	115				
		: :   :   :   :   :   :   :   :   :   :   :					
Dd	78	YTVMEWGEDIKAIVSKDEAVMLVNHOATGD--VCITLMCLDRLGLVVAQMMLMDHFXYT	135				
Oy	116	PLYGCYFAGHGSIYKKRSAKFENEKE---MRNKLQSVYAGTPMVIYIFPEGRTRYPEQT	172				
		: :   :   :   :   :   :   :   :   :					
Dd	136	N-FGIYSLVHGDFIRROGRSYRDQDLLLKHLLENNYSRDKRWIVLPEEGFLR---K	190				
Oy	173	VLSAQAFAAORGVLVKHVLTPRKATKTHVAEDCM---KN-----YLDA-----	213				
		: :   :       :   :   :   :   :   :					
Dd	191	KRETSQAAPRAKKNNLPLEFTLVNLLPRSGATKIILNALVAQQKNGSPAGDAKELDSKSLQ	250				
Oy	214	-IYDVTVVY 221					
		:					
Dd	251	WIDPTTIAY 259					
RESULT 9							
PLSC_YEAST	PLSC_YEAST	STANDARD;	PRT;	303 AA.			
ID AC	P33333:						
Dt	01-FEB-1994	(Rel. 28, Created)					
Dt	01-FEB-1994	(Rel. 28, Last sequence update)					
Dt	15-DEC-1998	(Rel. 37, Last annotation update)					

DE Probable 1-acyl-sn-glycerol-3-phosphate acyltransferase (EC 2.3.1.51)  
DE (1-AGP acyltransferase) (1-AGPAT) (lysophosphatidic acid  
DE acyltransferase) (LPAT).  
GN SL1 OR YDL052C.  
OS Saccharomyces cerevisiae (Baker's yeast).  
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;  
OC Saccharomycetales; Saccharomycetaceae; Saccharomyces.  
OX NCBI\_TaxID=4932;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=94012814; PubMed=8408076;  
RA Nagiec M.M., Wells G.B., Lester R.L., Dickson R.C.;  
RT "A suppressor gene that enables Saccharomyces cerevisiae to grow  
RT without making sphingolipids encodes a protein that resembles an  
RT Escherichia coli fatty acyltransferase."  
RL J. Biol. Chem. 268:22156-22163(1993).  
RN [2]  
RP SEQUENCE FROM N.A.  
RA Urestarazu L.A., Andre B., Vissers S.;  
RL Submitted (JUL-1996) to the EMBL/GenBank/DBJ databases.  
RN [3]  
RP SEQUENCE OF 182-303 FROM N.A.  
RA Bioecker H., Brandt P.;  
RL Submitted (JUL-1996) to the EMBL/GenBank/DBJ databases.  
CC -1- FUNCTION: MAY BE AN ACYLTRANSFERASE WITH AN ALTERED SUBSTRATE  
CC SPECIFICITY THAT ENABLES IT TO USE A C-26-COA IN PLACE OF THE  
CC C-16 OR C-18-COAS USED BY THE WILD TYPE PROTEIN.  
CC -1- CATALYTIC ACTIVITY: ACYL-CoA + 1-acyl-sn-glycerol 3-phosphate -  
CC CoA + 1,2-diacyl-sn-glycerol 3-phosphate.  
CC -1- PATHWAY: FIRST STEP IN DE NOVO PHOSPHOLIPID BIOSYNTHESIS.  
CC -1- SUBCELLULAR LOCATION: Integral membrane protein (potential).  
CC -1- SIMILARITY: BELONGS TO THE 1-ACYL-SN-GLYCEROL-3-PHOSPHATE  
CC ACYLTRANSFERASE FAMILY.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL: L13282; AAA16514.1; -;  
DR EMBL: Z74100; CAA98614.1; -;  
DR PIR: A48600; A48600.  
DR SGD: S0002210; SL1.  
DR InterPro: IPR002123; Acyltransferase.  
DR Pfam: PF01553; Acyltransferase; 1.  
KW Phospholipid biosynthesis; Transferase; Acyltransferase;  
KW Transmembrane.  
FT TRANSMEM 16 30 POTENTIAL.  
FT VARIANT 44 44 Q -> L (IN SUPPRESSOR ALLELE SL1-1).  
SQ SEQUENCE 303 AA; 33887 MW; 36ECB8C2659655EB CRC64;  
-----  
Query Match 10.0%; Score 120.5; DB 1; Length 303;  
Best Local Similarity 29.7%; Pred. No. 0.00067;  
Matches 44; Conservative 20; Mismatches 61; Indels 23; Gaps 6;  
-----  
OY 45 CVYOSVLFPEENTGYOILLYGDLPRKNKEIITLANHOSVDIVADILAIROMALGHV 104  
DB 50 CFYHVMKLM-----GLDVKVGGEENLAKPKYIMIANHOSLDTIFM-----LGR 94  
OY 105 -----RYVLEKGLKMLPLLYGCFYFAOHGIYKRSKFNKEKMKKLSQSYVDAGTPMYLVI 159  
DB 95 FPPCGTATKAKSLKLYVFLGFMALSGTYFLDRSKRQEAIDTLNKGLENYK-KRRALMV 153  
OY 160 FPEGSTR-YNPEQTKVLSASQAF-AAQRG 185  
DB 154 FPEGSTRYSTSLTMLPRKKAFAHQAOG 181  
-----  
RESULT 10

YIHG\_ECOLI  
ID YIHG\_ECOLI STANDARD; PRT; 310 AA.  
AC P31129;  
DT 01-OCT-1993 (Rel. 27, Created)  
DT 01-OCT-1993 (Rel. 27, Last sequence update)  
DT 16-OCT-2001 (Rel. 40, Last annotation update)  
DE Hypothetical protein yihg.  
GN YIHG OR B3862.  
OS Escherichia coli.  
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;  
OC Escherichia.  
OX NCBI\_TaxID=562;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=K12 / MG1655;  
RX MEDLINE=93347969; PubMed=8346018;  
RA Plunkett G., IIT, Burland V.D., Daniels D.L., Blattner F.R.;  
RT "Analysis of the Escherichia coli genome. IIT. DNA sequence of the  
RT region from 87.2 to 89.2 minutes."  
RL Nucleic Acids Res. 21:3391-3398(1993).  
CC -1- SIMILARITY: BELONGS TO THE 1-ACYL-SN-GLYCEROL-3-PHOSPHATE  
CC ACYLTRANSFERASE FAMILY.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL: L19201; AAB02997.1; -;  
DR EMBL: AE000461; AAC76860.1; -;  
DR PIR: S40808; S40808.  
DR EcGene: EG1833; Yihg.  
DR InterPro: IPR002123; Acyltransferase.  
DR Pfam: PF01553; Acyltransferase; 1.  
KW Hypothetical protein; Complete proteome.  
SQ SEQUENCE 310 AA; 36289 MW; 9F8E3F52EB0B186E CRC64;  
-----  
Query Match 10.0%; Score 120.5; DB 1; Length 310;  
Best Local Similarity 21.4%; Pred. No. 0.00067;  
Matches 50; Conservative 47; Mismatches 86; Indels 51; Gaps 10;  
-----  
OY 1 MRYLPSVVLGTAFTVLAAGVRLSAFLPAFYQALDRLCYVOSVWLFPEENTYG 60  
DB 43 VKLLLPVPV-----IMRKYSR-----CPFMKWCCEGLAVLHLN-PH 80  
OY 61 VOILLYGDLPKKNEN-IITLANHOSVDIVADILAIROMALGHV---RYVLEKGLKMLP 116  
DB 81 LOWEVHGLEGSKKNWILLICNHRSMWDIVLVCVLFKR-----HIPMKYFLKQOLAWPV 135  
OY 117 LYG--C-----YFAOHGIYKRSKFNKEKMKKLSQSYVDAGT-PMYLVIPPEGRYV 167  
DB 136 FLGLACWSLMDPEFKKRSRAVYLLRHPERRGKDVETTRRSCEKRLPHTTIVNVESSRFT 195  
OY 168 PEQGRKVLASQAFAQAGLAVLKHLVPRIKATGVAFADCKKNLDAIVDTVYV 221  
DB 196 QEK-----HQQTHSTFQNLPRPKAAGIALMALNVLGKQFDKLLNTLTCY 238  
-----  
RESULT 11  
PLSC\_LIMDO STANDARD; PRT; 281 AA.  
ID PLSC\_LIMDO  
AC 042870: 040120:  
DT 15-DEC-1998 (Rel. 37, Created)  
DT 15-DEC-1998 (Rel. 37, Last sequence update)  
DT 15-DEC-1998 (Rel. 37, Last annotation update)  
DE 1-acyl-sn-glycerol-3-phosphate acyltransferase (EC 2.3.1.51) (1-AGP  
DE acyltransferase) (1-AGPAT) (lysophosphatidic acid acyltransferase)  
DE (LPAT).  
GN PLSC.

```

OS Limnathes douglasii (Douglas's meadowfoam).
OC Eukaryota: Viridiplantae: Streptophyta: Embryophyta: Tracheophyta;
OC Spermatophyta: Magnoliophyta: eudicotyledons: core eudicots: Rosidae;
OC eucotsids II; Brassicales; Limnathaceae; Limnathes.
OX NCBI_taxid=28973;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=96028122; PubMed=7588719;
RA Hanke C., Wolter F.P., Coleman J., Peterrek G., Frentzen M.;
RT "A plant acyltransferase involved in triacylglycerol biosynthesis
RT complements an Escherichia coli sn-1-acylglycerol-3-phosphate
RT acyltransferase mutant."
RL Eur. J. Biochem. 232:806-810(1995).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=96046746; PubMed=7579178;
RA Brown A.P., Brough C.L., Kroon J., Stabas A.R.;
RT "Identification of a cDNA that encodes a 1-acyl-sn-glycerol-3-
RT phosphate acyltransferase from Limnathes douglasii."
RL Plant Mol. Biol. 29:267-278(1995).
CC -1- FUNCTION: CONVERTS LYSOPHOSPHATIDIC ACID (LPA) INTO PHOSPHATIDIC
CC ACID BY INCORPORATING ACYL MOIETY AT THE 2 POSITION. THIS ENZYME
CC USES ERUCOYL-COA AS AN ACYL DONOR.
CC -1- CATALYTIC ACTIVITY: ACYL-COA + 1-acyl-sn-glycerol 3-phosphate =
CC COA + 1,2-diacyl-sn-glycerol 3-phosphate.
CC -1- PATHWAY: SECOND STEP IN DE NOVO PHOSPHOLIPID BIOSYNTHESIS.
CC -1- SUBCELLULAR LOCATION: Integral membrane protein (Potential1).
CC -1- SIMILARITY: BELONGS TO THE 1-ACYL-SN-GLYCEROL-3-PHOSPHATE
CC ACYLTRANSFERASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as their content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-slb.ch/announce/
CC or send an email to license@isb-slb.ch).
CC -----
CC
CC EMBL; X83266; CAA58239.1; -.
CC DR EMBL; 246836; CAA86877.1; -.
CC DR InterPro: IPR001213; Acyltransferase.
CC KW Pfam: PF01553; Acyltransferase; 1.
CC DR Phospholipid biosynthesis; Transferase; Acyltransferase;
CC Transmembrane.
CC KM Transmem 40 60 POTENTIAL.
CC FT TRANSMEM 71 91 POTENTIAL.
CC FT TRANSMEM 110 130 POTENTIAL.
CC FT CONFLICT 46 46 I -> V (IN REF. 2).
CC FT CONFLICT 188 188 R -> G (IN REF. 2).
CC FT CONFLICT 262 262 V -> I (IN REF. 2).
CC FT CONFLICT 281 281 N -> K (IN REF. 2).
CC SQ SEQUENCE 281 AA; 31716 MM; 9C80BD9E492EE2A CRC64;

```

[illegible]



Query Match	9.6%	Score 116:	DB 1:	Length 281;
Best Local Similarity	24.2%	Pred. No. 0.0016:		
Matches 59;	Conservative 41;	Mismatches 86;	Indels 58;	Gaps 13.
QY	9	VLLGAPYYVLAMGWYRLLSAFLPARFYQALDRLY-CVYQSVMLEFFENYTGVOILLYG	67	
Db	46	IVLITA---VAMGL--IMVLILPPMYRIRIRGNLYGHILGLVIMLV---GIPLEIGG	95	
QY	68	DLPKKENIITLANQSVVDVMDIILAIROALNHVRVILKEGLKMYLPYGCYFRNHHG	127	
Db	96	S-EHTKKRAIYISINASPLDAEFVWMLA---PISTGVAKKEVL--WPLLQOLVTLAHN	149	
QY	128	IYVKSASFENEKEMHNKIQSVYDAGTP--MYLVIPEEGTRNYPEOTKVLASQAQ---AA	182	
Db	150	IRIDRS---NPAATIQSKKEAVRVTENLNSLIMPEGRSG--DGRLLPFKKGFVHLAL	204	
QY	183	QRCGLAVLKHVLTPIKATHVAF-----DCKNKYLDATY	215	
Db	205	QSHRPVYMITL---GTHLAWRKCTFPRVVPITVKKLPINTDDWTVDKIDYVKKMTH	260	
QY	216	DVTV	219	
Db	261	DIYV	264	
RESULT 14				
PLSC_MYCPN				
ID PLSC_MYCPN	STANDARD:	PRT:	266 AA.	
AC p75479;				
DT 01-NOV-1997 (Rel. 35, Created)				
DT 01-NOV-1997 (Rel. 35, last sequence update)				

Query Match	9.5%	Score 114	DB 1	Length 266
Best Local Similarity	25.5%	Pred. No. 0.0022		
Matches 56	Conservative 42	Mismatches 74	Indels 48	Gaps 13
QY	1	MRYL-LPSVYLIGTAPTVLWAGVWRLLSAFLPAIFYOALDD-----RLXCVYSMWLF	53	
Db	13	LRFLQLSLVLYV--LPYEVL-----MLISLSIAKNESIDENPPPEIRPKKVRVLSLF	65	
QY	54	FFENTGYQIILLGDLPRN--KENITLIYANHOSTYDMLVADILAIRNALGHRVYLKEG	111	
Db	66	LY--IKGVKVIIVN--PENVPKKALVLVYANKHNSMLDPLILKAFQKTEGVPPLTFEIAIE	121	
QY	112	LK--WL-----PLYGCFYFQHGIGYVKKRSARENEKEMRKLL--OSYVDAGTPMYLVIPRE	162	
Db	122	LQDPLWLFKIMLIDCVFD-----RKNLEKMAASLEQOOQIINQSTA--LCVEPE	169	
QY	163	GTRYNPEQTKVLSASQAFAGRGGLAVLKHVLTLPRIKATHV	202	
Db	170	GTR-----VLSRQIGEFKSGALKAVYANVPVPLTIV	202	
RESULT 15				
VENV_FQMPV				
ID	VENV_FQMPV	STANDARD	PRT	377 AA.
AC	P36316			
DT	01-JUN-1994 (Rel. 29, Created)			
DT	01-JUN-1994 (Rel. 29, Last sequence update)			
DE	16-OCT-2001 (Rel. 40, Last annotation update)			
DE	Major envelope protein (p43 kDa protein) (p43k).			
GN	FPV108.			
AS	Fowlpox virus (FPV).			



```

OC Viruses; dsDNA viruses, no RNA stage; Poxviridae; Chordopoxvirinae;
OC Avipoxvirus.
OX NCBI_TaxID:10261;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE-93079881; PubMed-1333124;
RA Calvert J.G., Ogawa R., Yanagida N., Nazerian K.;
RT "Identification and functional analysis of the fowlpox virus homolog
RT of the vaccinia virus p37k major envelope antigen gene.";
RL Virology 191:783-792(1992).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE-20193820; PubMed-10729156;
RA Afonso C.L., Tulman E.R., Lu Z., Zsak L., Kutish G.F., Rock D.L.;
RT "The genome of fowlpox virus.";
RL J. Virol. 74:3815-3831(2000).
CC -!- FUNCTION: THIS PROTEIN IS THE MAJOR ANTIGEN ON THE ENVELOPE OF
CC EXTRACELLULAR VIRUS.
CC -!- SIMILARITY: TO VACCINIA VIRUS 37 kDa ENVELOPE PROTEIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC
DR EMBL; M88587; AAA43819.1; -
DR EMBL; M88588; AAA47186.1; -
DR EMBL; AF198100; AAF4452.1; -
DR PIR; AA4216; AA4216.
DR InterPro; IPR001736; PLD.
DR Pfam; PF00614; PLDC; 2.
DR SMART; SM00155; PLDC; 2.
DR Transmembrane; Late protein.
SQ SEQUENCE 377 AA; 43021 MW; 5F466B7C97F9B86 CRC64;

```

Query Match 9.2%; Score 111; DB 1; Length 377;  
Best Local Similarity 20.9%; Pred. No. 0.0065;  
Matches 51; Conservative 53; Mismatches 80; Indels 60; Gaps 13;

```

QY 1 MRYLLPSVVLGTAPYVYVLAAGVWRLLSAFLPARFYQALDDRLX---CY---QSMVLF 53
DB 128 LHFYIGSASLGNALTYITNNGIYSNNSLAMDLYFRSLDYKTISSKKCLEFTPMATKYH 187
QY 54 FFENYGVQIILYGDLPKKEN-----IYLANHQSIVDWIVADIL----- 94
DB 188 FFRNNGI---FFSDSPENHVGKRRFEDDCVHYITDAKSTIDLAVSLPFRKTRKDSI 244
QY 95 ---AIFQNALGHVRYVLEKGLKWLPLYGCFYAGHGSIYVRSRK-FNEKEMRNKLQSYVD 150
DB 245 VYVPIIKDAL--IRAVLEKGVKLRVLG--FWKKTDOVISKASIKSLNELSV-----DHID 295
QY 151 AGTPMWLVTFPESTRYNP-EQTKVLSASQAFAQRLAVLKHYLTPRIKATH-----VA 203
DB 296 ISTKYR--FPVNSKVDDINNSKMTIDGRYA-----HVMYTNLDGSHRNHAEVS 344
QY 204 FDCM 207
DB 345 FNCM 348

```

Search completed: August 28, 2002, 11:17:24  
Job time: 548 sec

**THIS PAGE BLANK (USPTO)**

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 28, 2002, 11:16:56 ; Search time 65.95 Seconds

(without alignments)  
598.072 Million cell updates/sec

Title: US-09-853-526-70

Perfect score: 1203

Sequence: 1 MRYLLPSVVLGTAPRYVLA.....NYLDATYDVYVEGKDDG 228

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 562222 seqs, 172994929 residues

Total number of hits satisfying chosen parameters: 562222

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :  
1: SPREMBL\_19:\*  
2: sp\_archaea:\*  
3: sp\_bacteria:\*  
4: sp\_fungi:\*  
5: sp\_human:\*  
6: sp\_invertebrate:\*  
7: sp\_mammal:\*  
8: sp\_mhc:\*  
9: sp\_organelle:\*  
10: sp\_plant:\*  
11: sp\_rotent:\*  
12: sp\_virus:\*  
13: sp\_vertebrate:\*  
14: sp\_unclassified:\*  
15: sp\_virus:\*  
16: sp\_bacteriap:\*  
17: sp\_archaeap:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1203	100.0	364	4 Q9B0G4	Q9B0G4 homo sapien
2	1200	99.8	353	4 Q9N0G2	Q9N0G2 homo sapien
3	1040	86.5	354	11 Q9D1E8	Q9D1E8 mus musculu
4	338	28.1	386	5 Q9S1R2	Q9S1R2 caenorhabdi
5	312	25.9	373	10 Q9SSH0	Q9SSH0 arabidopsis
6	312	25.9	393	10 Q9C9P8	Q9C9P8 arabidopsis
7	291	24.2	375	10 Q9L1M4	Q9L1M4 arabidopsis
8	269	22.4	374	10 Q41745	Q41745 zea mays (m
9	240	20.0	306	10 Q9SDN3	Q9SDN3 prunus dulc
10	238	19.8	377	10 Q40119	Q40119 limanthes
11	235.5	19.6	390	10 Q9XFM4	Q9XFM4 brassica na
12	213.5	17.7	376	12 Q9SYC8	Q9SYC8 arabidopsis
13	210	17.5	386	5 Q9VVS1	Q9VVS1 drosophila
14	203	16.9	380	5 Q9VVA9	Q9VVA9 drosophila
15	200	16.6	344	5 Q16526	Q16526 caenorhabdi
16	197.5	16.4	378	11 Q924S1	Q924S1 rattus norv

17	178.5	14.8	316	11 Q9DB84	Q9DB84 mus musculu
18	172.5	14.3	314	11 Q9D517	Q9D517 mus musculu
19	171	14.2	311	10 Q93J17	Q93J17 brassica na
20	168.5	14.0	310	10 Q9SVX9	Q9SVX9 arabidopsis
21	164	13.6	399	5 Q23087	Q23087 caenorhabdi
22	161.5	13.4	350	3 Q94361	Q94361 schizosacch
23	142	11.8	359	10 Q9FF57	Q9FF57 arabidopsis
24	141.5	11.8	426	5 Q01882	Q01882 caenorhabdi
25	141.5	11.8	428	5 Q9SR03	Q9SR03 caenorhabdi
26	141	11.7	295	16 Q91657	Q91657 pseudomonas
27	134	11.1	439	5 Q20800	Q20800 caenorhabdi
28	133.5	11.1	370	11 Q91YX5	Q91YX5 mus musculu
29	116	9.6	281	10 Q9SD02	Q9SD02 limanthes
30	113	9.4	247	16 Q9X219	Q9X219 thermotoga
31	112.5	9.4	248	2 Q30849	Q30849 salmonella
32	112	9.3	258	16 Q9HW50	Q9HW50 pseudomonas
33	112	9.3	304	16 Q91546	Q91546 pseudomonas
34	109.5	9.1	235	2 Q9KX14	Q9KX14 rhodobacter
35	108	9.0	261	16 Q9KPE3	Q9KPE3 vibrio chol
36	107	8.9	241	16 Q9CPE2	Q9CPE2 pasteurella
37	105.5	8.8	249	2 Q9EY25	Q9EY25 xanthomonas
38	105.5	8.8	262	16 Q98G17	Q98G17 rhizobium 1
39	105	8.7	260	16 Q92MG4	Q92MG4 rhizobium m
40	104	8.6	344	10 Q9L1Y4	Q9L1Y4 brassica na
41	102.5	8.5	291	3 Q9Y7C5	Q9Y7C5 emericella
42	102	8.5	212	10 Q9M0A2	Q9M0A2 arabidopsis
43	101.5	8.4	292	16 Q9K1N9	Q9K1N9 rhizobium m
44	101	8.4	323	16 Q9KAG4	Q9KAG4 bacillus ba
45	99	8.2	243	16 Q9ZD75	Q9ZD75 rickettsia

## ALIGNMENTS

RESULT 1  
Q9B0G4 PRELIMINARY: PRT; 364 AA.  
ID Q9B0G4  
AC Q9B0G4:  
DT 01-JUN-2001 (TREMBLrel. 17, Created)  
DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)  
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)  
DE HYPOTHEICAL 42.1 KDA PROTEIN (LYSOPHOSPHATIDIC ACID  
DE ACYLTANSFERASE-EPSILON) (EC 2.3.1.51).  
GN DKFPP61C22 OR LPAAT-E.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxId=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=AMYGDAIA;  
RX MEDLINE=21154917; PubMed=11230166;  
RA Wiemann S., Weil B., Wellenreuther R., Gassenhuber J., Glassl S.,  
RA Ansoerge W., Boecker M., Bloeker H., Bauersachs S., Blum H.,  
RA Lauber J., Duesterhoeft A., Beyer A., Koehrer K., Strack N.,  
RA Mewes H.W., Ottenwaelder B., Obermaier B., Tampe J., Heubner D.,  
RA Wandt R., Korn B., Klein M., Rouska A.,  
RA "Towards a Catalog of Human Genes and Proteins: Sequencing and  
RT Analysis of 500 Novel Complete Protein Coding Human cDNAs";  
RT Genome Res. 11:422-435(2001).  
[2]  
RP SEQUENCE FROM N.A.  
RA Leung D.W.,  
RT "Cloning and expression of LPAAT-epsilon";  
RT Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AL136587; CAB6522.1; -  
DR EMBL; AF375789; AAK54809.1; -  
DR InterPro: IPR002123; Acyltransferase.  
DR Pfam: PF01553; Acyltransferase; 1.  
KW Hypothetical protein; transferase; Acyltransferase.  
SQ SEQUENCE 364 AA: 42072 MW: 90A0F87FC7C78081 CRC64;





[illegible]

DT 01-MAY-2000 (TREMBLrel. 13, Created)  
DT 01-MAY-2000 (TREMBLrel. 13, last sequence update)  
DT 01-DEC-2001 (TREMBLrel. 19, last annotation update)  
DE 1-ACYL-SN-GLYCEROL-3-PHOSPHATE ACYLTRANSFERASE.  
DE Prunus dulcis (Almond) (*Prunus amygdalus*).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliopsida; euclidyales; core eudicots; Rosidae  
OC eucosids I; Rosales; Rosaceae; Amygdaloidae; Prunus.  
OX NCBI\_TaxId=3755;  
RN [1]\_\_\_\_\_  
RP SEQUENCE FROM N.A.  
RA Campalans A., Pages M., Messguier R.;  
RT "Identification of differentially expressed genes during dehydration

RESULT	12
09SYC8	
ID	09SYC8
AC	PRELIMINARY;
DT	01-MAY-2000 (TREMBLrel. 13, Created)
DT	01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT	01-JUN-2001 (TREMBLrel. 17, Last annotation update)
DE	PUTATIVE ACYL-COA:1-ACYLGlycerol-3-PHOSPHATE ACYLTRANSFERASE.
GN	FLIML5.12.
OS	Arabidopsis thaliana (Mouse-ear cress).
OC	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC	Spermatophytes; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC	eumustoids; II; Brassicales; Brassicaceae; Arabidopsids.
OX	NCBI_TaxID=3702;
RN	
RM	[1]
RP	SEQUENCE FROM N.A.
RA	Federerpiet N.A., Palm C.J., Conway A.B., Conn L., Hansen N.F.,
RA	Altafi H., Araujo R., Huizar L., Rowley D., Buehler E., Dunn P.,
RA	Gonzalez A., Kremenetskaia I., Kim C., Lenz C., Li J., Liu S.,
RA	Lunos S., Schwartz J., Shinn P., Toriumi M., Vysotskaia V.S.,





RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,  
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,  
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,  
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,  
RA Foslter C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,  
RA Glöckel A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,  
RA Harris N.L., Harvey D., Helman T.J., Hernandez J.R., Houck J.,  
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Idegawa C.,  
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,  
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,  
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,  
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,  
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,  
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,  
RA Nelson D.R., Nelson K.A., Nixon K., Nuskern D.R., Paclebo J.M.,  
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,  
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,  
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,  
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,  
RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,  
RA Wang Z.-Y., Massarman D.A., Weinstein G.M., Weissenbach J.,  
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,  
RA Ye J., Yen R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,  
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,  
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;  
RT "The genome sequence of *Drosophila melanogaster*.";  
RL Science 287:2185-2195(2000).  
DR EMBL: AE003527; AAF49473.1; -  
DR FlyBase: FBgn0036622; CG4753.  
DR InterPro: IPR002123; Acyltransferase.  
DR Pfam: PF01553; Acyltransferase; 1.  
SQ SEQUENCE 380 AA; 43482 MW; EA240947F566A841 CRC64;

Query Match 16.9%; Score 203; DB 5; Length 380;  
Best Local Similarity 27.5%; Pred. No. 1.2e-11;  
Matches 56; Conservative 47; Mismatches 79; Indels 22; Gaps 5;  
QY 25 RLISAPLPARFYQALDRLKCYVOSWLYFFENYTGQIILLXGDLPRN-----KENIIT 78  
DB 43 RSLAVYHHSFYCI---LVCV-----AEWYAGSKLKVYID-PQDEQKFFQKEHGL 89  
QY 79 LANHSTVDIVADILAIRONALGHVRYLKEGLKWLPLXCYFAQHGITYKRSKAFNE 138  
DB 90 LMNHTYEIDMLTAMITDKLNGIGCTAKYAKKMLRYVPVGLGWMMAFELDRNFEKDK 149  
QY 139 KEMRNKIQSYVDAGTPMYLIPFGRTYRNPQGTQVLSASQAFPAQRLAVLKHVLPRIK 198  
DB 150 VVIKTQKEVFSYDPDPVWLLNAGTRFTPAKHKL---SVKFAERGLPLKLHLLIPRTK 206  
QY 199 ATHTVAFDCKNYLDATYDVTVYE 222  
DB 207 GFTTSLPTMRGICPAIYDINIAFR 230

RESULT 15  
ID 016526 PRELIMINARY; PRT; 344 AA.  
AC 016526;  
DT 01-JAN-1998 (TREMBlrel. 05, Created)  
DT 01-JAN-1998 (TREMBlrel. 05, Last sequence update)  
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)  
DE T05H4.1 PROTEIN.  
GN T05H4.1.  
OS Caenorhabditis elegans.  
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;  
OC Rhabditidae; Pelodierinae; Caenorhabditis.  
OX NCBI\_Taxid=6239;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-BRISTOL N2;  
RX MEDLINE=94150718; PubMed=7906398;  
WILSON R., Ainscough R., Anderson K., Baynes C., Berks M.,

RA Bonfield J., Burton J., Connell M., Copsey T., Cooper J., Coulson A.,  
RA Craxton M., Dear S., Du Z., Durbin R., Favello A., Fulton L.,  
RA Gardner A., Green P., Hawkins T., Hillier L., Jier M., Johnston L.,  
RA Jones M., Kershaw J., Kirsten J., Laister N., Latreille P.,  
RA Lightning J., Lloyd C., McMurray A., Mortimore B., O'Callaghan M.,  
RA Parsons J., Percy C., Rifkin L., Roopra A., Saunders D., Showman R.,  
RA Shalton N., Smith A., Sonhammer E., Staden R., Sulston J.,  
RA Thierly-Mieg J., Thomas K., Vaudin M., Vaughan K., Waterston R.,  
RA Watson A., Weinstein L., Wilkinson-Sproat J., Wohldman P.;  
RT "2.2 Mb of contiguous nucleotide sequence from chromosome III of *C.*  
RT *elegans*.";  
RL Nature 368:32-38(1994).  
RL Nature 368:32-38(1994).  
RP SEQUENCE FROM N.A.  
RC STRAIN-BRISTOL N2;  
RP SEQUENCE FROM N.A.  
RT "The sequence of *C. elegans* cosmid T05H4.";  
RT Submitted (AUG-1997) to the EMBL/GenBank/DBJ databases.  
RL [3]  
RP SEQUENCE FROM N.A.  
RC STRAIN-BRISTOL N2;  
RA Waterston R.;  
RL Submitted (JUL-1997) to the EMBL/GenBank/DBJ databases.  
DR EMBL: AF016452; AAB66008.1; -  
DR InterPro: IPR002123; Acyltransferase.  
DR Pfam: PF01553; Acyltransferase; 1.  
SQ SEQUENCE 344 AA; 40070 MW; ABD8CF47038BE48 CRC64;

Query Match 16.6%; Score 200; DB 5; Length 344;  
Best Local Similarity 28.7%; Pred. No. 2e-11;  
Matches 66; Conservative 35; Mismatches 93; Indels 36; Gaps 10;  
QY 10 LIGTA---PYYLAWGWRLSLAPLPARFYQALDRLKCYVOSWLYFFENYTGQIILL 65  
DB 18 LIGTVFLPLPLDLAW-----FAP-KLMRCADRLVGFWLPFCSLIEWGVNFRV 68  
QY 66 YGDLPRKKNIIYLANHSTVDIVADILAIRONALGHV-----RYLKEGLKWLPL 117  
DB 69 TGDILIERDEPAIILMNRHTKLDWLFs-----WNAIKKMPWLLTTERKISLAPLKKIP- 121  
QY 118 YGCVFAQHGIGIYKRSKAF-NEKEMRNKIQSYVDAGTPMY-LVIFEGRTYRNPQGTQVLS 175  
DB 122 -GAGWMASSGSYIFLDRNFENDKPVLEIRIYKYSGSEKKYQIILLFAGTDKGERARL- 178  
QY 176 ASQAFPAQRLAVLKHVLPRIKATHVAFDCKR--NLDATYDVTVYEG 223  
DB 179 -SDAFADKNGLPREYVLIHPRTTGFKFLMELMKKENYIKYVDLTAYSG 227

Search completed: August 28, 2002, 11:16:57  
Job time: 561 sec



GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 28, 2002, 11:14:17 ; Search time 30.82 Seconds  
(without alignments)  
279.761 Million cell updates/sec

Title: US-09-853-526-4

Perfect score: 353

Sequence: 1 MRYLLPSVVLGTAFTYVLA.....YVNTWIVGTLGLMTWTKA 353

Scoring table: OLIGO

Gapop 60.0 , Gapext 60.0

Searched: 231628 seqs, 24425594 residues

Word size : 0

Total number of hits satisfying chosen parameters: 231628

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database : Issued\_Patents\_AA:\*

1: /cgn2\_6/ptodata/2/1aa/5A.COMB.pep:\*  
2: /cgn2\_6/ptodata/2/1aa/5B.COMB.pep:\*  
3: /cgn2\_6/ptodata/2/1aa/6A.COMB.pep:\*  
4: /cgn2\_6/ptodata/2/1aa/6B.COMB.pep:\*  
5: /cgn2\_6/ptodata/2/1aa/CTUS.COMB.pep:\*  
6: /cgn2\_6/ptodata/2/1aa/Dackfile1.pep:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	353	100.0	353	2	US-08-996-306-4
2	353	100.0	353	4	US-09-338-907-4
3	353	100.0	353	4	US-09-218-207-4
4	353	100.0	364	2	US-08-996-306-5
5	353	100.0	364	4	US-09-338-907-5
6	353	100.0	364	4	US-09-218-207-5
7	230	65.2	291	4	US-09-338-907-127
8	230	65.2	291	4	US-09-218-207-127
9	228	64.6	228	4	US-09-338-907-70
10	228	64.6	228	4	US-09-218-207-70
11	199	56.4	261	4	US-09-338-907-128
12	199	56.4	261	4	US-09-218-207-128
13	185	52.4	185	4	US-09-338-907-136
14	185	52.4	185	4	US-09-218-207-136
15	185	52.4	315	4	US-09-338-907-134
16	185	52.4	315	4	US-09-218-207-134
17	184	52.1	300	4	US-09-338-907-135
18	184	52.1	300	4	US-09-218-207-135
19	154	43.6	182	4	US-09-338-907-133
20	154	43.6	182	4	US-09-218-207-133
21	116	32.9	238	4	US-09-338-907-126
22	116	32.9	238	4	US-09-218-207-126
23	85	24.1	97	4	US-09-338-907-132
24	85	24.1	97	4	US-09-218-207-132
25	62	17.6	66	4	US-09-338-907-131
26	62	17.6	66	4	US-09-218-207-131
27	62	17.6	68	4	US-09-338-907-130

28	62	17.6	68	4	US-09-218-207-130	Sequence 130, App
29	62	17.6	77	4	US-09-338-907-125	Sequence 125, App
30	62	17.6	77	4	US-09-218-207-125	Sequence 125, App
31	62	17.6	90	4	US-09-338-907-129	Sequence 129, App
32	62	17.6	90	4	US-09-218-207-129	Sequence 129, App
33	32	9.1	354	4	US-09-338-907-74	Sequence 74, App
34	32	9.1	354	4	US-09-218-207-74	Sequence 74, App
35	9	2.5	9	2	US-08-996-306-15	Sequence 15, App
36	9	2.5	9	4	US-09-338-907-15	Sequence 15, App
37	9	2.5	9	4	US-09-218-207-15	Sequence 15, App
38	8	2.3	9	2	US-08-996-306-18	Sequence 18, App
39	8	2.3	9	4	US-09-338-907-18	Sequence 18, App
40	8	2.3	9	4	US-09-218-207-18	Sequence 18, App
41	7	2.0	12	1	US-08-224-625-1	Sequence 1, App
42	7	2.0	12	2	US-08-254-404-1	Sequence 1, App
43	7	2.0	12	2	US-08-327-451E-1	Sequence 1, App
44	7	2.0	12	2	US-08-458-109-1	Sequence 1, App
45	7	2.0	12	3	US-08-231-196-1	Sequence 1, App

## ALIGNMENTS

RESULT 1  
US-08-996-306-4  
; Sequence 4, Application US/08996306  
; Patent No. 5945522  
; GENERAL INFORMATION:  
; APPLICANT: Cohen, Daniel  
; APPLICANT: Chumakov, Ilya  
; APPLICANT: Blumenfeld, Marta  
; APPLICANT: Bougueret, Lydie  
; TITLE OF INVENTION: Prostate cancer gene  
; NUMBER OF SEQUENCES: 68  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Knobbe, Martens, Olson & Bear  
; STREET: 501 West Broadway  
; CITY: San Diego  
; STATE: California  
; COUNTRY: USA  
; ZIP: 92101-3505  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy Disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: Win95  
; SOFTWARE: Word  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/996,306  
; FILING DATE:  
; CLASSIFICATION:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Israelson, Ned A.  
; REGISTRATION NUMBER: 29,655  
; REFERENCE/DOCKET NUMBER: GENSET.018A  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (619) 235-8550  
; TELEFAX: (619) 235-0176  
; INFORMATION FOR SEQ ID NO: 4:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 353 amino acids  
; TYPE: AMINO ACID  
; STRANDEDNESS: SINGLE  
; TOPOLOGY: LINEAR  
; MOLECULE TYPE: PROTEIN  
; ORIGINAL SOURCE:  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: potential Transmembrane helix  
; LOCATION: 1..33  
; IDENTIFICATION METHOD: Rao and Argos method  
; FEATURE:  
; NAME/KEY: potential Transmembrane helix  
; LOCATION: 4..20

IDENTIFICATION METHOD: Klein, Kanehisa and Delisi method  
FEATURE:  
NAME/KEY: potential Transmembrane helix  
LOCATION: 4..24  
IDENTIFICATION METHOD: Eisenberg, Schwarz, Komarony  
IDENTIFICATION METHOD: and Wall method  
FEATURE:  
NAME/KEY: potential N-myristoylation site  
LOCATION: 12  
IDENTIFICATION METHOD: prosite match  
FEATURE:  
NAME/KEY: potential Transmembrane helix  
LOCATION: 50..70  
IDENTIFICATION METHOD: Eisenberg, Schwarz, Komarony and Wall method  
FEATURE:  
NAME/KEY: potential N-glycosylation site  
LOCATION: 57  
IDENTIFICATION METHOD: prosite match  
FEATURE:  
NAME/KEY: potential Transmembrane helix  
LOCATION: 76..96  
IDENTIFICATION METHOD: Eisenberg, Schwarz, Komarony and Wall method  
FEATURE:  
NAME/KEY: potential Tyrosine kinase phosphorylation site  
LOCATION: 78  
IDENTIFICATION METHOD: prosite match  
FEATURE:  
NAME/KEY: potential Casein kinase II phosphorylation site  
LOCATION: 84  
IDENTIFICATION METHOD: prosite match  
FEATURE:  
NAME/KEY: potential Leucine zipper pattern  
LOCATION: 94..115  
IDENTIFICATION METHOD: prosite match  
FEATURE:  
NAME/KEY: potential N-myristoylation site  
LOCATION: 119  
IDENTIFICATION METHOD: prosite match  
FEATURE:  
NAME/KEY: potential Protein kinase C phosphorylation site  
LOCATION: 133  
IDENTIFICATION METHOD: prosite match  
FEATURE:  
NAME/KEY: potential Casein kinase II phosphorylation site  
LOCATION: 147  
IDENTIFICATION METHOD: prosite match  
FEATURE:  
NAME/KEY: potential Protein kinase C phosphorylation site  
LOCATION: 194  
IDENTIFICATION METHOD: prosite match  
FEATURE:  
NAME/KEY: potential Tyrosine kinase phosphorylation site  
LOCATION: 215  
IDENTIFICATION METHOD: prosite match  
FEATURE:  
NAME/KEY: potential Tyrosine sulfatation site  
LOCATION: 221  
IDENTIFICATION METHOD: prosite match  
FEATURE:  
NAME/KEY: potential cAMP- and cGMP-dependent protein kinase phosphorylation site  
LOCATION: 223  
IDENTIFICATION METHOD: prosite match  
FEATURE:  
NAME/KEY: potential Casein kinase II phosphorylation site  
LOCATION: 235  
IDENTIFICATION METHOD: prosite match  
FEATURE:  
NAME/KEY: potential Protein kinase C phosphorylation site  
LOCATION: 306  
IDENTIFICATION METHOD: prosite match  
FEATURE:  
NAME/KEY: potential Transmembrane helix  
LOCATION: 310..330

IDENTIFICATION METHOD: Eisenberg, Schwarz, Komarony and Wall method  
FEATURE:  
NAME/KEY: potential N-myristoylation site  
LOCATION: 319  
IDENTIFICATION METHOD: prosite match  
FEATURE:  
NAME/KEY: potential N-myristoylation site  
LOCATION: 323  
IDENTIFICATION METHOD: prosite match  
FEATURE:  
NAME/KEY: potential Amladation site  
LOCATION: 329  
IDENTIFICATION METHOD: prosite match  
FEATURE:  
NAME/KEY: potential Transmembrane helix  
LOCATION: 333..353  
IDENTIFICATION METHOD: Eisenberg, Schwarz, Komarony and Wall method  
FEATURE:  
NAME/KEY: potential N-myristoylation site  
LOCATION: 341  
IDENTIFICATION METHOD: prosite match  
FEATURE:  
NAME/KEY: potential Protein kinase C phosphorylation site  
LOCATION: 350  
IDENTIFICATION METHOD: prosite match  
US-08-996-306-4  
Query Match 100.0%; Score 353; DB 2; Length 353;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 353; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 MRYLLPSVVLGTAFTVLAAGVWRLLSAFLPARFYQALDRLCYVQSMVLEFFENYTG 60  
DB 1 MRYLLPSVVLGTAFTVLAAGVWRLLSAFLPARFYQALDRLCYVQSMVLEFFENYTG 60  
QY 61 VOILLYDLDPRKNENIYYLANHOSTVDVIADIAIQNALGHVRYLKSGKMLPLYGC 120  
DB 61 VOILLYDLDPRKNENIYYLANHOSTVDVIADIAIQNALGHVRYLKSGKMLPLYGC 120  
QY 121 YFAOHGSIYVRSKAKFENEKEMRNKLOSYVDAGPMYLVIFPEGRYNPEOTKYLSASQAF 180  
DB 121 YFAOHGSIYVRSKAKFENEKEMRNKLOSYVDAGPMYLVIFPEGRYNPEOTKYLSASQAF 180  
QY 181 AAORGGLAVLKHVLPRIKATIHVAFDCMKNYLDAIYDVTVVYEGKDDGQRRSEPTWTEFL 240  
DB 181 AAORGGLAVLKHVLPRIKATIHVAFDCMKNYLDAIYDVTVVYEGKDDGQRRSEPTWTEFL 240  
QY 241 CKECPRKIHIDRIDKADVPEDEHMRMWLHERFEIKDKMLIEFYESPDEPRRRKRPFGKS 300  
DB 241 CKECPRKIHIDRIDKADVPEDEHMRMWLHERFEIKDKMLIEFYESPDEPRRRKRPFGKS 300  
QY 301 VNSKLSIKKTLPSMLILSGLTAGMLMTDAGRKLVTNWTIVGTLGCLWVITKA 353  
DB 301 VNSKLSIKKTLPSMLILSGLTAGMLMTDAGRKLVTNWTIVGTLGCLWVITKA 353  
RESULT 2  
US-09-338-907-4  
; Sequence 4, Application US/09338907  
; Patent No. 6265546  
; GENERAL INFORMATION:  
; APPLICANT: Cohen, Daniel  
; APPLICANT: Blumenfeld, Marla  
; APPLICANT: Ilya, Chumakov  
; APPLICANT: Bouguetelret, Lydie  
; TITLE OF INVENTION: PROSTATE CANCER GENE  
; FILE REFERENCE: GENSET.18CPICP  
; CURRENT APPLICATION NUMBER: US/09/338,907  
; CURRENT FILING DATE: 1999-06-23  
; EARLIER APPLICATION NUMBER: 08/996,306  
; EARLIER FILING DATE: 1997-12-22  
; EARLIER APPLICATION NUMBER: 60/099,658

EARLIER FILING DATE: 1998-09-09  
EARLIER APPLICATION NUMBER: 09/218,207  
EARLIER FILING DATE: 1998-12-22  
NUMBER OF SEQ ID NOS: 578  
SOFTWARE: Patent.pm  
SEQ ID NO 4  
LENGTH: 353  
TYPE: PRT  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: HELIX  
LOCATION: 1..33  
OTHER INFORMATION: Rao and Argos identification method, potential helix  
FEATURE:  
NAME/KEY: HELIX  
LOCATION: 4..20  
OTHER INFORMATION: Klein, Kanehisa and Delisi identification method, potential  
FEATURE:  
NAME/KEY: HELIX  
LOCATION: 4..24  
OTHER INFORMATION: Eisenberg, Schwarz, Komarony, Wall identification method, F  
FEATURE:  
NAME/KEY: MYRISTATE  
LOCATION: 12..16  
OTHER INFORMATION: Prosite match  
FEATURE:  
NAME/KEY: HELIX  
LOCATION: 50..70  
OTHER INFORMATION: Eisenberg, Schwarz, Komarony, Wall identification method, F  
FEATURE:  
NAME/KEY: CARBOHYD  
LOCATION: 57..59  
OTHER INFORMATION: Prosite match  
FEATURE:  
NAME/KEY: HELIX  
LOCATION: 76..96  
OTHER INFORMATION: Eisenberg, Schwarz, Komarony, Wall identification method, F  
FEATURE:  
NAME/KEY: PHOSPHORYLATION  
LOCATION: 78  
OTHER INFORMATION: potential Tyrosine kinase site, Prosite match  
FEATURE:  
NAME/KEY: PHOSPHORYLATION  
LOCATION: 84  
OTHER INFORMATION: potential caseine kinase II site, Prosite match  
FEATURE:  
NAME/KEY: SITE  
LOCATION: 94..115  
OTHER INFORMATION: potential leucine zipper site, Prosite match  
FEATURE:  
NAME/KEY: MYRISTATE  
LOCATION: 119..123  
OTHER INFORMATION: potential site, Prosite match  
FEATURE:  
NAME/KEY: PHOSPHORYLATION  
LOCATION: 133  
OTHER INFORMATION: potential protein kinase C, Prosite match  
FEATURE:  
NAME/KEY: PHOSPHORYLATION  
LOCATION: 147  
OTHER INFORMATION: potential caseine kinase II site, Prosite match  
FEATURE:  
NAME/KEY: PHOSPHORYLATION  
LOCATION: 194  
OTHER INFORMATION: potential protein kinase C, Prosite match  
FEATURE:  
NAME/KEY: PHOSPHORYLATION  
LOCATION: 215  
OTHER INFORMATION: potential Tyrosine kinase site, Prosite match  
FEATURE:  
NAME/KEY: SULFATATION  
LOCATION: 221  
OTHER INFORMATION: Prosite match

FEATURE:  
NAME/KEY: PHOSPHORYLATION  
LOCATION: 233  
OTHER INFORMATION: potential cAMP and cGMP dependant protein kinase site,  
FEATURE:  
NAME/KEY: PHOSPHORYLATION  
LOCATION: 235  
OTHER INFORMATION: potential caseine kinase II site, Prosite match  
FEATURE:  
NAME/KEY: PHOSPHORYLATION  
LOCATION: 306  
OTHER INFORMATION: potential protein kinase C, Prosite match  
FEATURE:  
NAME/KEY: HELIX  
LOCATION: 310..330  
OTHER INFORMATION: Eisenberg, Schwarz, Komarony, Wall identification method,  
FEATURE:  
NAME/KEY: MYRISTATE  
LOCATION: 319..323  
OTHER INFORMATION: Prosite match  
FEATURE:  
NAME/KEY: MYRISTATE  
LOCATION: 323..327  
OTHER INFORMATION: Prosite match  
FEATURE:  
NAME/KEY: AMIDATION  
LOCATION: 329  
OTHER INFORMATION: Prosite match  
FEATURE:  
NAME/KEY: HELIX  
LOCATION: 333..353  
OTHER INFORMATION: Eisenberg, Schwarz, Komarony, Wall identification method,  
FEATURE:  
NAME/KEY: MYRISTATE  
LOCATION: 341..345  
OTHER INFORMATION: Prosite match  
FEATURE:  
NAME/KEY: PHOSPHORYLATION  
LOCATION: 350  
OTHER INFORMATION: potential protein kinase C, Prosite match  
US-09-338-907-4  
Query Match 100.0%; Score 353; DB 4; Length 353;  
Best Local Similarity 100.0%; Pred No. 0; Mismatches 0; Indels 0; Gaps 0;  
Matches 353; Conservative 0;  
QY 1 MRYLLPSVVLGTAPTYVLAMGVWRLLSAFLPARFYQALDDRLCYQSYVTFEFPENYTG 60  
DB 1 MRYLLPSVVLGTAPTYVLAMGVWRLLSAFLPARFYQALDDRLCYQSYVTFEFPENYTG 60  
QY 61 VOILLYGDLPKKNENIYLANHOSYVDWIVADILAIRONALGHVRYVLEGLKWLPLYGC 120  
DB 61 VOILLYGDLPKKNENIYLANHOSYVDWIVADILAIRONALGHVRYVLEGLKWLPLYGC 120  
QY 121 YFAOHGIIYKRSKAKNEKEMRNKLOSVDAGTPTVLVTFPESTRNPBOTKVLASQAF 180  
DB 121 YFAOHGIIYKRSKAKNEKEMRNKLOSVDAGTPTVLVTFPESTRNPBOTKVLASQAF 180  
QY 181 AAOGLAVLKHVLTPIRIKATHVAFDCMKNYLDAIYDVTVVYEGKDDGGORRESPTMTEFL 240  
DB 181 AAOGLAVLKHVLTPIRIKATHVAFDCMKNYLDAIYDVTVVYEGKDDGGORRESPTMTEFL 240  
QY 241 CKCEPKIHIDRIKDKVDPEEOEHMRWLHEFEIKDKMLIEFYSPDPERKRRPGKS 300  
DB 241 CKCEPKIHIDRIKDKVDPEEOEHMRWLHEFEIKDKMLIEFYSPDPERKRRPGKS 300  
QY 301 VNSKLSIKTLPRLMLISGLTAGMLMTDAGRKLYVTMTWITGLGLMTWITKA 353  
DB 301 VNSKLSIKTLPRLMLISGLTAGMLMTDAGRKLYVTMTWITGLGLMTWITKA 353  
RESULT 3

```
US-09-218-207-4
; Sequence 4, Application US/09218207
; Patent No. 6346381
; GENERAL INFORMATION:
; APPLICANT: Cohen, Daniel
; APPLICANT: Blumenfeld, Marta
; APPLICANT: Ilya, Chumakov
; APPLICANT: Bouqueloret, Lydie
; TITLE OF INVENTION: Prostate cancer gene
; FILE REFERENCE: GENSET.018CP1
; CURRENT APPLICATION NUMBER: US/09/218,207
; EARLIER FILING DATE: 1998-12-22
; EARLIER FILING DATE: 1997-12-22
; EARLIER APPLICATION NUMBER: 60/099,658
; EARLIER FILING DATE: 1998-09-09
; NUMBER OF SEQ ID NOS: 578
; SOFTWARE: Patent.pm
; SEQ ID NO 4
; LENGTH: 353
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: HELIX
; LOCATION: 1..33
; OTHER INFORMATION: Rao and Argos Identification method, potential helix
; FEATURE:
; NAME/KEY: HELIX
; LOCATION: 4..20
; OTHER INFORMATION: Klein, Kanehisa and Delisi Identification method, potential
; FEATURE:
; NAME/KEY: HELIX
; LOCATION: 4..24
; OTHER INFORMATION: Eisenberg, Schwarz, Komarony, Wall Identification method,
; FEATURE:
; NAME/KEY: MYRISTATE
; LOCATION: 12..16
; OTHER INFORMATION: Prositte match
; FEATURE:
; NAME/KEY: HELIX
; LOCATION: 50..70
; OTHER INFORMATION: Eisenberg, Schwarz, Komarony, Wall Identification method,
; FEATURE:
; NAME/KEY: CARBOHYD
; LOCATION: 57..59
; OTHER INFORMATION: Prositte match
; FEATURE:
; NAME/KEY: HELIX
; LOCATION: 76..96
; OTHER INFORMATION: Eisenberg, Schwarz, Komarony, Wall Identification method,
; FEATURE:
; NAME/KEY: PHOSPHORYLATION
; LOCATION: 78
; OTHER INFORMATION: potential Tyrosine kinase site, Prositte match
; FEATURE:
; NAME/KEY: PHOSPHORYLATION
; LOCATION: 84
; OTHER INFORMATION: potential caseine kinase II site, Prositte match
; FEATURE:
; NAME/KEY: SITE
; LOCATION: 94..115
; OTHER INFORMATION: potential Leucine zipper site, Prositte match
; FEATURE:
; NAME/KEY: MYRISTATE
; LOCATION: 119..123
; OTHER INFORMATION: potential site, Prositte match
; FEATURE:
; NAME/KEY: PHOSPHORYLATION
; LOCATION: 133
; OTHER INFORMATION: potential protein kinase C, Prositte match
; FEATURE:
; NAME/KEY: PHOSPHORYLATION
; LOCATION: 147
; OTHER INFORMATION: potential caseine kinase II site, Prositte match
; FEATURE:
; NAME/KEY: PHOSPHORYLATION
; LOCATION: 194
; OTHER INFORMATION: potential protein kinase C, Prositte match
; FEATURE:
; NAME/KEY: PHOSPHORYLATION
; LOCATION: 215
; OTHER INFORMATION: potential Tyrosine kinase site, Prositte match
; FEATURE:
; NAME/KEY: SULFATATION
; LOCATION: 221
; OTHER INFORMATION: Prositte match
; FEATURE:
; NAME/KEY: PHOSPHORYLATION
; LOCATION: 233
; OTHER INFORMATION: potential CAMP and cGMP dependant protein kinase site,
; FEATURE:
; NAME/KEY: PHOSPHORYLATION
; LOCATION: 235
; OTHER INFORMATION: potential caseine kinase II site, Prositte match
; FEATURE:
; NAME/KEY: PHOSPHORYLATION
; LOCATION: 306
; OTHER INFORMATION: potential protein kinase C, Prositte match
; FEATURE:
; NAME/KEY: HELIX
; LOCATION: 310..330
; OTHER INFORMATION: Eisenberg, Schwarz, Komarony, Wall Identification method,
; FEATURE:
; NAME/KEY: MYRISTATE
; LOCATION: 319..323
; OTHER INFORMATION: Prositte match
; FEATURE:
; NAME/KEY: MYRISTATE
; LOCATION: 323..327
; OTHER INFORMATION: Prositte match
; FEATURE:
; NAME/KEY: AMIDATION
; LOCATION: 329
; OTHER INFORMATION: Prositte match
; FEATURE:
; NAME/KEY: HELIX
; LOCATION: 333..353
; OTHER INFORMATION: Eisenberg, Schwarz, Komarony, Wall Identification method,
; FEATURE:
; NAME/KEY: MYRISTATE
; LOCATION: 341..345
; OTHER INFORMATION: Prositte match
; FEATURE:
; NAME/KEY: PHOSPHORYLATION
; LOCATION: 350
; OTHER INFORMATION: potential protein kinase C, Prositte match
; OTHER INFORMATION: potential protein kinase C, Prositte match
US-09-218-207-4
Query Match 100.0%; Score 353; DB 4; Length 353;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 353; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MRYLLPSVVLGTAFTYVLAAGWRLLSAPLPAFTYALDRLCYQSVNLFEEFNTG 60
Db 1 MRYLLPSVVLGTAFTYVLAAGWRLLSAPLPAFTYALDRLCYQSVNLFEEFNTG 60
QY 61 VOILLYGDLPKNKENTIIYLANHSTVMIYADILAIROMNLGHRVYLKXGLKWLPLYGC 120
Db 61 VOILLYGDLPKNKENTIIYLANHSTVMIYADILAIROMNLGHRVYLKXGLKWLPLYGC 120
QY 121 YFAQHGSIYVRSKAFNEKEMRNKLSQSYVDAGTPMYLVIFPEGTRVNPEDQTKVLSASQAF 180
Db 121 YFAQHGSIYVRSKAFNEKEMRNKLSQSYVDAGTPMYLVIFPEGTRVNPEDQTKVLSASQAF 180
QY 181 AAGRLAVLKHVLTPIKATVHAFDCMKNTLDALIVDTVYVEGKDDGGQRRESPTMTTEFL 240
```

Db 181 AAORGLAVLKHVLTFRKATHVAFDCMKNYLDAYDTVYVEKDDGGGQRESPTMEFL 240  
QY 241 CKECPRIHIDRIDKKDVEEBOEHMRMLHEFEIKDKMLIEFYESSPDERRRKRPFGKS 300  
Db 241 CKECPRIHIDRIDKKDVEEBOEHMRMLHEFEIKDKMLIEFYESSPDERRRKRPFGKS 300  
QY 301 VNSKLSIKKTLPMSMLISGLTAGMLMTDAGRKLVTWNTWYIGTLGCLMWTIKA 353  
Db 301 VNSKLSIKKTLPMSMLISGLTAGMLMTDAGRKLVTWNTWYIGTLGCLMWTIKA 353

## RESULT 4

US-08-996-306-5  
; Sequence 5, Application US/08996306  
; Patent No. 5945522  
; GENERAL INFORMATION:  
; APPLICANT: Cohen, Daniel  
; APPLICANT: Chumakov, Ilya  
; APPLICANT: Blumenfeld, Marta  
; APPLICANT: Bougueleret, Lydie  
; TITLE OF INVENTION: Prostate cancer gene  
; NUMBER OF SEQUENCES: 68  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Knobbe, Martens, Olson & Bear  
; STREET: 501 West Broadway  
; CITY: San Diego  
; STATE: California  
; COUNTRY: USA  
; ZIP: 92101-3505  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy Disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: Win95  
; SOFTWARE: Word  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/996,306  
; FILING DATE:  
; CLASSIFICATION:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Istaelsen, Ned A.  
; REGISTRATION NUMBER: 29,655  
; REFERENCE/DOCKET NUMBER: GENSET.018A  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (619) 235-8550  
; TELEFAX: (619) 235-0176  
; INFORMATION FOR SEQ ID NO: 5:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 364 amino acids  
; TYPE: AMINO ACID  
; STRANDEDNESS: SINGLE  
; TOPOLOGY: LINEAR  
; MOLECULE TYPE: PROTEIN  
; ORIGINAL SOURCE:  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: potential protein  
; LOCATION: 1..364  
; US-08-996-306-5

Query Match 100.0%; Score 353; DB 2; Length 364;

Best Local Similarity 100.0%; Pred. No. 0; Mismatches 0; Indels 0; Gaps 0;

Matches 353; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MRYLLPSVVLGTAPRYVLAMGWRLLSAFLPARFYQALDDRLCYVYOSWVLEFFENYTG 60  
Db 12 MRYLLPSVVLGTAPRYVLAMGWRLLSAFLPARFYQALDDRLCYVYOSWVLEFFENYTG 71  
QY 61 VOILLYGDLPKKNENIYLANHOSYDWMIVADILAIRONALGHVRYVLEKGLKMLPLYGC 120  
Db 72 VOILLYGDLPKKNENIYLANHOSYDWMIVADILAIRONALGHVRYVLEKGLKMLPLYGC 131

QY 121 YFAOHGIYVYKRSKAKFENEKEMRNKLOSVDAGTPTMLVIFPESTRNPEQTKVLSAQAF 180  
Db 132 YFAOHGIYVYKRSKAKFENEKEMRNKLOSVDAGTPTMLVIFPESTRNPEQTKVLSAQAF 191  
QY 181 AAORGLAVLKHVLTFRKATHVAFDCMKNYLDAYDTVYVEKDDGGGQRESPTMEFL 240  
Db 192 AAORGLAVLKHVLTFRKATHVAFDCMKNYLDAYDTVYVEKDDGGGQRESPTMEFL 251  
QY 241 CKECPRIHIDRIDKKDVEEBOEHMRMLHEFEIKDKMLIEFYESSPDERRRKRPFGKS 300  
Db 252 CKECPRIHIDRIDKKDVEEBOEHMRMLHEFEIKDKMLIEFYESSPDERRRKRPFGKS 311  
QY 301 VNSKLSIKKTLPMSMLISGLTAGMLMTDAGRKLVTWNTWYIGTLGCLMWTIKA 353  
Db 312 VNSKLSIKKTLPMSMLISGLTAGMLMTDAGRKLVTWNTWYIGTLGCLMWTIKA 364

## RESULT 5

US-09-338-907-5  
; Sequence 5, Application US/09338907  
; Patent No. 6265546  
; GENERAL INFORMATION:  
; APPLICANT: Cohen, Daniel  
; APPLICANT: Blumenfeld, Marta  
; APPLICANT: Ilya, Chumakov  
; APPLICANT: Bougueleret, Lydie  
; TITLE OF INVENTION: PROSTATE CANCER GENE  
; FILE REFERENCE: GENSET.18CPICP  
; CURRENT APPLICATION NUMBER: US/09/338,907  
; CURRENT FILING DATE: 1999-06-23  
; EARLIER APPLICATION NUMBER: 08/996,306  
; EARLIER FILING DATE: 1997-12-22  
; EARLIER APPLICATION NUMBER: 60/099,658  
; EARLIER FILING DATE: 1998-09-09  
; EARLIER APPLICATION NUMBER: 09/218,207  
; EARLIER FILING DATE: 1998-12-22  
; NUMBER OF SEQ ID NOS: 578  
; SOFTWARE: Patent.pm  
; SEQ ID NO 5  
; LENGTH: 364  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; US-09-338-907-5

Query Match 100.0%; Score 353; DB 4; Length 364;

Best Local Similarity 100.0%; Pred. No. 0; Mismatches 0; Indels 0; Gaps 0;

Matches 353; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MRYLLPSVVLGTAPRYVLAMGWRLLSAFLPARFYQALDDRLCYVYOSWVLEFFENYTG 60  
Db 12 MRYLLPSVVLGTAPRYVLAMGWRLLSAFLPARFYQALDDRLCYVYOSWVLEFFENYTG 71  
QY 61 VOILLYGDLPKKNENIYLANHOSYDWMIVADILAIRONALGHVRYVLEKGLKMLPLYGC 120  
Db 72 VOILLYGDLPKKNENIYLANHOSYDWMIVADILAIRONALGHVRYVLEKGLKMLPLYGC 131  
QY 121 YFAOHGIYVYKRSKAKFENEKEMRNKLOSVDAGTPTMLVIFPESTRNPEQTKVLSAQAF 180  
Db 132 YFAOHGIYVYKRSKAKFENEKEMRNKLOSVDAGTPTMLVIFPESTRNPEQTKVLSAQAF 191  
QY 181 AAORGLAVLKHVLTFRKATHVAFDCMKNYLDAYDTVYVEKDDGGGQRESPTMEFL 240  
Db 192 AAORGLAVLKHVLTFRKATHVAFDCMKNYLDAYDTVYVEKDDGGGQRESPTMEFL 251  
QY 241 CKECPRIHIDRIDKKDVEEBOEHMRMLHEFEIKDKMLIEFYESSPDERRRKRPFGKS 300  
Db 252 CKECPRIHIDRIDKKDVEEBOEHMRMLHEFEIKDKMLIEFYESSPDERRRKRPFGKS 311  
QY 301 VNSKLSIKKTLPMSMLISGLTAGMLMTDAGRKLVTWNTWYIGTLGCLMWTIKA 353  
Db 312 VNSKLSIKKTLPMSMLISGLTAGMLMTDAGRKLVTWNTWYIGTLGCLMWTIKA 364

RESULT 6  
US-09-218-207-5  
; Sequence 5, Application US/09218207  
; Patent No. 6346381  
; GENERAL INFORMATION:  
; APPLICANT: Cohen, Daniel  
; APPLICANT: Blumenfeld, Marta  
; APPLICANT: Ilya, Chumakov  
; APPLICANT: Bougueleret, Lydie  
; TITLE OF INVENTION: Prostate cancer gene  
; FILE REFERENCE: GENSET.018CP1  
; CURRENT APPLICATION NUMBER: US/09/218,207  
; CURRENT FILING DATE: 1998-12-22  
; EARLIER APPLICATION NUMBER: 08/996,306  
; EARLIER FILING DATE: 1997-12-22  
; EARLIER APPLICATION NUMBER: 60/099,658  
; EARLIER FILING DATE: 1998-09-09  
; NUMBER OF SEQ ID NOS: 578  
; SOFTWARE: Patent.pm  
; SEQ ID NO 5  
; LENGTH: 364  
; TYPE: PR1  
; ORGANISM: Homo sapiens  
US-09-218-207-5

Query Match 100.0%; Score 353; DB 4; Length 364;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 353; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MRYLLPSVLLGTAPTYVLAAGVWRLSAPFLPARFYQALDRLCYQSVNLYFFFEYNTG 60  
DB 12 MRYLLPSVLLGTAPTYVLAAGVWRLSAPFLPARFYQALDRLCYQSVNLYFFFEYNTG 71  
QY 61 VOILLYGDLPRKNENITYLANHSTVMIYADILAIQNALGHRVYLKEGLKPLRYGC 120  
DB 72 VOILLYGDLPRKNENITYLANHSTVMIYADILAIQNALGHRVYLKEGLKPLRYGC 131  
QY 121 YFAOHGSIYVRSKAFNEKEMRNKLOSVDAGTPMYLVIFPEGRYVPEOTKVLASQAF 180  
DB 132 YFAOHGSIYVRSKAFNEKEMRNKLOSVDAGTPMYLVIFPEGRYVPEOTKVLASQAF 191  
QY 181 AAOGLAVLKHVLTLPRIKATHVAFDCMKNYLDIYDVTVYEGKDGQORRESPTMEFL 240  
DB 192 AAOGLAVLKHVLTLPRIKATHVAFDCMKNYLDIYDVTVYEGKDGQORRESPTMEFL 251  
QY 241 CKCEPKIHIDRIKDVPEDEHMRWLHEREIKDKMLIEFYESPDPERRRRPFGKS 300  
DB 252 CKCEPKIHIDRIKDVPEDEHMRWLHEREIKDKMLIEFYESPDPERRRRPFGKS 311  
QY 301 VNSKLSIKTLPMSMLISGLTAGMLMTDAGRKLXVNTWYIGTLGCLMWTIKA 353  
DB 312 VNSKLSIKTLPMSMLISGLTAGMLMTDAGRKLXVNTWYIGTLGCLMWTIKA 364

RESULT 7  
US-09-338-907-127  
; Sequence 127, Application US/09338907  
; Patent No. 6265546  
; GENERAL INFORMATION:  
; APPLICANT: Cohen, Daniel  
; APPLICANT: Blumenfeld, Marta  
; APPLICANT: Ilya, Chumakov  
; APPLICANT: Bougueleret, Lydie  
; TITLE OF INVENTION: PROSTATE CANCER GENE  
; FILE REFERENCE: GENSET.18CP1CP  
; CURRENT APPLICATION NUMBER: US/09/338,907  
; CURRENT FILING DATE: 1999-06-23  
; EARLIER APPLICATION NUMBER: 08/996,306  
; EARLIER FILING DATE: 1997-12-22  
; EARLIER APPLICATION NUMBER: 60/099,658  
; EARLIER FILING DATE: 1998-09-09

; EARLIER APPLICATION NUMBER: 09/218,207  
; EARLIER FILING DATE: 1998-12-22  
; NUMBER OF SEQ ID NOS: 578  
; SOFTWARE: Patent.pm  
; SEQ ID NO 127  
; LENGTH: 291  
; TYPE: PR1  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: SITE  
; LOCATION: 98..103  
; OTHER INFORMATION: Box II  
; NAME/KEY: SITE  
; LOCATION: 149..157  
; OTHER INFORMATION: Box III  
US-09-338-907-127

Query Match 65.2%; Score 230; DB 4; Length 291;  
Best Local Similarity 100.0%; Pred. No. 6e-223;  
Matches 230; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 124 OHGSIYVRSKAFNEKEMRNKLOSVDAGTPMYLVIFPEGRYVPEOTKVLASQAF 183  
DB 62 OHGSIYVRSKAFNEKEMRNKLOSVDAGTPMYLVIFPEGRYVPEOTKVLASQAF 121  
QY 184 RGLAVLKHVLTLPRIKATHVAFDCMKNYLDIYDVTVYEGKDGQORRESPTMEFLCKE 243  
DB 122 RGLAVLKHVLTLPRIKATHVAFDCMKNYLDIYDVTVYEGKDGQORRESPTMEFLCKE 181  
QY 244 CPKIHIDRIKDVPEDEHMRWLHEREIKDKMLIEFYESPDPERRRRPFGKS 303  
DB 182 CPKIHIDRIKDVPEDEHMRWLHEREIKDKMLIEFYESPDPERRRRPFGKS 241  
QY 304 KLSIKTLPMSMLISGLTAGMLMTDAGRKLXVNTWYIGTLGCLMWTIKA 353  
DB 242 KLSIKTLPMSMLISGLTAGMLMTDAGRKLXVNTWYIGTLGCLMWTIKA 291

RESULT 8  
US-09-218-207-127  
; Sequence 127, Application US/09218207  
; Patent No. 6346381  
; GENERAL INFORMATION:  
; APPLICANT: Cohen, Daniel  
; APPLICANT: Blumenfeld, Marta  
; APPLICANT: Ilya, Chumakov  
; APPLICANT: Bougueleret, Lydie  
; TITLE OF INVENTION: Prostate cancer gene  
; FILE REFERENCE: GENSET.018CP1  
; CURRENT APPLICATION NUMBER: US/09/218,207  
; CURRENT FILING DATE: 1998-12-22  
; EARLIER APPLICATION NUMBER: 08/996,306  
; EARLIER FILING DATE: 1997-12-22  
; EARLIER APPLICATION NUMBER: 60/099,658  
; EARLIER FILING DATE: 1998-09-09  
; NUMBER OF SEQ ID NOS: 578  
; SOFTWARE: Patent.pm  
; SEQ ID NO 127  
; LENGTH: 291  
; TYPE: PR1  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: SITE  
; LOCATION: 98..103  
; OTHER INFORMATION: Box II  
; NAME/KEY: SITE  
; LOCATION: 149..157  
; OTHER INFORMATION: Box III  
US-09-218-207-127





OTHER INFORMATION: Box II  
FEATURE:  
NAME/KEY: SITE  
LOCATION: 119..127  
OTHER INFORMATION: Box III  
US-09-338-907-128

Query Match 56.4%; Score 199; DB 4; Length 261;  
Best Local Similarity 100.0%; Pred. No. 7.2e-192;  
Matches 199; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 155 MYLVIPEEGTRYNPEQTKVLSASQAFNAQGLAVLKHVLPRIKATHVAFDCMKNYLDAI 214  
|  
DB 63 MYLVIPEEGTRYNPEQTKVLSASQAFNAQGLAVLKHVLPRIKATHVAFDCMKNYLDAI 122  
QY 215 YDVTVVEGDDGQRESPTMTFEFLCKECPKIHIDRIDKKDVPPEQEHMRRLHERF 274  
|  
DB 123 YDVTVVEGDDGQRESPTMTFEFLCKECPKIHIDRIDKKDVPPEQEHMRRLHERF 182  
QY 275 EIKDKMLIEFESPDPRRRKFPCKSVNSKLSIKTLPMSLISGLTAGMLMTDAGRKL 334  
|  
DB 183 EIKDKMLIEFESPDPRRRKFPCKSVNSKLSIKTLPMSLISGLTAGMLMTDAGRKL 242  
QY 335 VNTWIYGTLLGCLMTWTIKA 353  
|  
DB 243 VNTWIYGTLLGCLMTWTIKA 261

RESULT 12  
US-09-218-207-128  
Sequence 128, Application US/09218207  
Patent No. 6346381  
GENERAL INFORMATION:  
APPLICANT: Cohen, Daniel  
APPLICANT: Blumenfeld, Marta  
APPLICANT: Ilya, Chumakov  
APPLICANT: Bougueleret, Lydie  
TITLE OF INVENTION: Prostate cancer gene  
FILE REFERENCE: GENSET.018CPI  
CURRENT APPLICATION NUMBER: US/09/218,207  
EARLIER FILING DATE: 1998-12-22  
EARLIER APPLICATION NUMBER: 08/996,306  
EARLIER FILING DATE: 1997-12-22  
EARLIER APPLICATION NUMBER: 60/099,658  
NUMBER OF SEQ ID NOS: 578  
SOFTWARE: Patent.pm  
SEQ ID NO 128  
LENGTH: 261  
TYPE: PRT  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: SITE  
LOCATION: 68..73  
OTHER INFORMATION: Box II  
FEATURE:  
NAME/KEY: SITE  
LOCATION: 119..127  
OTHER INFORMATION: Box III  
US-09-218-207-128

Query Match 56.4%; Score 199; DB 4; Length 261;  
Best Local Similarity 100.0%; Pred. No. 7.2e-192;  
Matches 199; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 155 MYLVIPEEGTRYNPEQTKVLSASQAFNAQGLAVLKHVLPRIKATHVAFDCMKNYLDAI 214  
|  
DB 63 MYLVIPEEGTRYNPEQTKVLSASQAFNAQGLAVLKHVLPRIKATHVAFDCMKNYLDAI 122  
QY 215 YDVTVVEGDDGQRESPTMTFEFLCKECPKIHIDRIDKKDVPPEQEHMRRLHERF 274  
|  
DB 123 YDVTVVEGDDGQRESPTMTFEFLCKECPKIHIDRIDKKDVPPEQEHMRRLHERF 182  
QY 275 EIKDKMLIEFESPDPRRRKFPCKSVNSKLSIKTLPMSLISGLTAGMLMTDAGRKL 334  
|  
DB 183 EIKDKMLIEFESPDPRRRKFPCKSVNSKLSIKTLPMSLISGLTAGMLMTDAGRKL 242  
QY 335 VNTWIYGTLLGCLMTWTIKA 353  
|  
DB 243 VNTWIYGTLLGCLMTWTIKA 261

DB 123 YDVTVVEGDDGQRESPTMTFEFLCKECPKIHIDRIDKKDVPPEQEHMRRLHERF 182  
QY 275 EIKDKMLIEFESPDPRRRKFPCKSVNSKLSIKTLPMSLISGLTAGMLMTDAGRKL 334  
|  
DB 183 EIKDKMLIEFESPDPRRRKFPCKSVNSKLSIKTLPMSLISGLTAGMLMTDAGRKL 242  
QY 335 VNTWIYGTLLGCLMTWTIKA 353  
|  
DB 243 VNTWIYGTLLGCLMTWTIKA 261

RESULT 13  
US-09-338-907-136  
Sequence 136, Application US/09338907  
Patent No. 6265546  
GENERAL INFORMATION:  
APPLICANT: Cohen, Daniel  
APPLICANT: Blumenfeld, Marta  
APPLICANT: Ilya, Chumakov  
APPLICANT: Bougueleret, Lydie  
TITLE OF INVENTION: PROSTATE CANCER GENE  
FILE REFERENCE: GENSET.18CPI  
CURRENT APPLICATION NUMBER: US/09/338,907  
CURRENT FILING DATE: 1999-06-23  
EARLIER APPLICATION NUMBER: 08/996,306  
EARLIER FILING DATE: 1997-12-22  
EARLIER APPLICATION NUMBER: 60/099,658  
EARLIER FILING DATE: 1998-09-09  
EARLIER APPLICATION NUMBER: 09/218,207  
EARLIER FILING DATE: 1998-12-22  
NUMBER OF SEQ ID NOS: 578  
SOFTWARE: Patent.pm  
SEQ ID NO 136  
LENGTH: 185  
TYPE: PRT  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: SITE  
LOCATION: 81..83  
OTHER INFORMATION: Box I  
FEATURE:  
NAME/KEY: SITE  
LOCATION: 160..165  
OTHER INFORMATION: Box II  
US-09-338-907-136

Query Match 52.4%; Score 185; DB 4; Length 185;  
Best Local Similarity 100.0%; Pred. No. 5.9e-178;  
Matches 185; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MRYLLPSVVLGTAFTVYLVAMGVRILSAFLPARFYQALDRLXCVQSWLFFEFNYTG 60  
|  
DB 1 MRYLLPSVVLGTAFTVYLVAMGVRILSAFLPARFYQALDRLXCVQSWLFFEFNYTG 60  
QY 61 VOILLVGLDLPKRNENIYLANHGSTVDMIVADILAIRQNALGHVRYLKGKMLPLYGC 120  
|  
DB 61 VOILLVGLDLPKRNENIYLANHGSTVDMIVADILAIRQNALGHVRYLKGKMLPLYGC 120  
QY 121 YFAOHGIIYVRSKAKFNEKEMRNKLOSVDAGTPMYLVIPEEGTRYNPEQTKVLSASQAF 180  
|  
DB 121 YFAOHGIIYVRSKAKFNEKEMRNKLOSVDAGTPMYLVIPEEGTRYNPEQTKVLSASQAF 180  
QY 181 AAORG 185  
|  
DB 181 AAORG 185

RESULT 14  
US-09-218-207-136  
Sequence 136, Application US/09218207  
Patent No. 6346381  
GENERAL INFORMATION:

```
; APPLICANT: Cohen, Daniel
; APPLICANT: Blumenfeld, Marta
; APPLICANT: Ilya, Chumakov
; APPLICANT: Bougueleret, Lydie
; TITLE OF INVENTION: Prostate cancer gene
; FILE REFERENCE: GENSET.018CPI
; CURRENT APPLICATION NUMBER: US/09/218,207
; CURRENT FILING DATE: 1998-12-22
; EARLIER APPLICATION NUMBER: 08/996,306
; EARLIER FILING DATE: 1997-12-22
; EARLIER APPLICATION NUMBER: 60/099,658
; EARLIER FILING DATE: 1998-09-09
; NUMBER OF SEQ ID NOS: 578
; SOFTWARE: Patent.pm
; SEQ ID NO 136
; LENGTH: 185
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: 81..83
; OTHER INFORMATION: Box I
; FEATURE:
; NAME/KEY: SITE
; LOCATION: 160..165
; OTHER INFORMATION: Box II
US-09-218-207-136
```

```
Query Match          52.4%; Score 185; DB 4; Length 185;
Best Local Similarity 100.0%; Pred. No. 5.9e-178;
Matches 185; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY 1 MRYLLPSVYLLGTAPTYVLAMGVMRLSAFLPARFYQALDDRLXCYYQSMVLEFFENYTG 60
    |||||||
DB 1 MRYLLPSVYLLGTAPTYVLAMGVMRLSAFLPARFYQALDDRLXCYYQSMVLEFFENYTG 60
OY 61 VOILLYGDLPRKNKENTITLANHOSTVDWIVADILAIRONALGHVRYLKEGLKWLPLYGC 120
    |||||||
DB 61 VOILLYGDLPRKNKENTITLANHOSTVDWIVADILAIRONALGHVRYLKEGLKWLPLYGC 120
OY 121 YFAOHGIGTYVKSRAKFNEMKRNKLOSVDAGTPMYLVIFPEGSTRYNPEQTKVLSASQAF 180
    |||||||
DB 121 YFAOHGIGTYVKSRAKFNEMKRNKLOSVDAGTPMYLVIFPEGSTRYNPEQTKVLSASQAF 180
OY 181 AAORG 185
    |||||
DB 181 AAORG 185
```

```
RESULT 15
US-09-338-907-134
; Sequence 134, Application US/09338907
; Patent No. 6265546
; GENERAL INFORMATION:
; APPLICANT: Cohen, Daniel
; APPLICANT: Blumenfeld, Marta
; APPLICANT: Ilya, Chumakov
; APPLICANT: Bougueleret, Lydie
; TITLE OF INVENTION: PROSTATE CANCER GENE
; FILE REFERENCE: GENSET.18CPICP
; CURRENT APPLICATION NUMBER: US/09/338,907
; CURRENT FILING DATE: 1999-06-23
; EARLIER APPLICATION NUMBER: 08/996,306
; EARLIER FILING DATE: 1997-12-22
; EARLIER APPLICATION NUMBER: 60/099,658
; EARLIER FILING DATE: 1998-09-09
; EARLIER APPLICATION NUMBER: 09/218,207
; EARLIER FILING DATE: 1998-12-22
; NUMBER OF SEQ ID NOS: 578
; SOFTWARE: Patent.pm
; SEQ ID NO 134
; LENGTH: 315
```

```
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: 81..83
; OTHER INFORMATION: Box I
; FEATURE:
; NAME/KEY: SITE
; LOCATION: 160..165
; OTHER INFORMATION: Box II
US-09-338-907-134
```

```
Query Match          52.4%; Score 185; DB 4; Length 315;
Best Local Similarity 100.0%; Pred. No. 9.7e-178;
Matches 185; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY 1 MRYLLPSVYLLGTAPTYVLAMGVMRLSAFLPARFYQALDDRLXCYYQSMVLEFFENYTG 60
    |||||||
DB 1 MRYLLPSVYLLGTAPTYVLAMGVMRLSAFLPARFYQALDDRLXCYYQSMVLEFFENYTG 60
OY 61 VOILLYGDLPRKNKENTITLANHOSTVDWIVADILAIRONALGHVRYLKEGLKWLPLYGC 120
    |||||||
DB 61 VOILLYGDLPRKNKENTITLANHOSTVDWIVADILAIRONALGHVRYLKEGLKWLPLYGC 120
OY 121 YFAOHGIGTYVKSRAKFNEMKRNKLOSVDAGTPMYLVIFPEGSTRYNPEQTKVLSASQAF 180
    |||||||
DB 121 YFAOHGIGTYVKSRAKFNEMKRNKLOSVDAGTPMYLVIFPEGSTRYNPEQTKVLSASQAF 180
OY 181 AAORG 185
    |||||
DB 181 AAORG 185
```

```
Search completed: August 28, 2002, 11:19:21
Job time: 304 sec
```

---

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 28, 2002, 11:17:02 ; Search time 305.3 Seconds  
(without alignments)  
406.974 Million cell updates/sec

Title: US-09-853-526-4

Perfect score: 353  
Sequence: 1 MRLLPSVLLGTAPTYVLA.....YVNTWIVGTLGLMTWTKA 353

Scoring table: GATCOP 60.0 , Gapext 60.0

Searched: 3502263 segs, 351980561 residues

Word size : 0

Total number of hits satisfying chosen parameters: 3502263

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database : Pending\_Patents\_AA\_Main:\*

- 1: /cgn2\_6/ptodata/2/paa/PCrus.COMB.pep.\*
- 2: /cgn2\_6/ptodata/2/paa/US06.COMB.pep.\*
- 3: /cgn2\_6/ptodata/2/paa/US07.COMB.pep.\*
- 4: /cgn2\_6/ptodata/2/paa/US080.COMB.pep.\*
- 5: /cgn2\_6/ptodata/2/paa/US081.COMB.pep.\*
- 6: /cgn2\_6/ptodata/2/paa/US082.COMB.pep.\*
- 7: /cgn2\_6/ptodata/2/paa/US083.COMB.pep.\*
- 8: /cgn2\_6/ptodata/2/paa/US084.COMB.pep.\*
- 9: /cgn2\_6/ptodata/2/paa/US085.COMB.pep.\*
- 10: /cgn2\_6/ptodata/2/paa/US086.COMB.pep.\*
- 11: /cgn2\_6/ptodata/2/paa/US087.COMB.pep.\*
- 12: /cgn2\_6/ptodata/2/paa/US088.COMB.pep.\*
- 13: /cgn2\_6/ptodata/2/paa/US089.COMB.pep.\*
- 14: /cgn2\_6/ptodata/2/paa/US090.COMB.pep.\*
- 15: /cgn2\_6/ptodata/2/paa/US091.COMB.pep.\*
- 16: /cgn2\_6/ptodata/2/paa/US092.COMB.pep.\*
- 17: /cgn2\_6/ptodata/2/paa/US093.COMB.pep.\*
- 18: /cgn2\_6/ptodata/2/paa/US094.COMB.pep.\*
- 19: /cgn2\_6/ptodata/2/paa/US095.COMB.pep.\*
- 20: /cgn2\_6/ptodata/2/paa/US096.COMB.pep.\*
- 21: /cgn2\_6/ptodata/2/paa/US097.COMB.pep.\*
- 22: /cgn2\_6/ptodata/2/paa/US098.COMB.pep.\*
- 23: /cgn2\_6/ptodata/2/paa/US099.COMB.pep.\*
- 24: /cgn2\_6/ptodata/2/paa/US100.COMB.pep.\*
- 25: /cgn2\_6/ptodata/2/paa/US101.COMB.pep.\*
- 26: /cgn2\_6/ptodata/2/paa/US60.COMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
1	353	100.0	353	1	PCT-US01-01431-59
2	353	100.0	353	1	PCT-US01-01431-79
3	353	100.0	353	1	PCT-US01-11988-1474
4	353	100.0	353	1	PCT-US01-11988-1475
5	353	100.0	353	22	US-09-853-245-1474
6	353	100.0	353	22	US-09-853-245-1475
7	353	100.0	353	22	US-09-853-526-4

8	353	100.0	353	23	US-09-901-484A-4	Sequence 4, Appl1
9	353	100.0	353	23	US-09-915-582-59	Sequence 59, Appl1
10	353	100.0	353	23	US-09-915-582-79	Sequence 79, Appl1
11	353	100.0	353	26	US-60-099-658-4	Sequence 4, Appl1
12	353	100.0	353	18	US-09-436-919-1	Sequence 4, Appl1
13	353	100.0	364	22	US-09-817-910-7	Sequence 7, Appl1
14	353	100.0	364	22	US-09-853-526-5	Sequence 5, Appl1
15	353	100.0	364	23	US-09-901-484A-5	Sequence 5, Appl1
16	353	100.0	364	26	US-60-099-658-5	Sequence 5, Appl1
17	290	82.2	353	18	US-09-488-725A-2736	Sequence 2736, Ap
18	269	76.2	269	1	PCT-US01-01327-47	Sequence 47, Appl1
19	252	71.4	372	18	US-09-488-725A-6308	Sequence 6308, Ap
20	248	70.3	450	26	US-60-245-222-110	Sequence 110, App
21	230	65.2	291	22	US-09-853-526-127	Sequence 127, App
22	230	65.2	291	23	US-09-901-484A-127	Sequence 127, App
23	228	64.6	228	22	US-09-853-526-70	Sequence 70, Appl1
24	228	64.6	228	23	US-09-901-484A-70	Sequence 70, Appl1
25	228	64.6	228	26	US-60-099-658-70	Sequence 128, App
26	199	56.4	261	22	US-09-853-526-128	Sequence 128, App
27	199	56.4	261	23	US-09-901-484A-128	Sequence 128, App
28	185	52.4	185	22	US-09-853-526-136	Sequence 136, App
29	185	52.4	185	23	US-09-901-484A-136	Sequence 136, App
30	185	52.4	315	22	US-09-853-526-134	Sequence 134, App
31	185	52.4	315	23	US-09-901-484A-134	Sequence 134, App
32	184	52.1	300	22	US-09-853-526-135	Sequence 135, App
33	184	52.1	300	23	US-09-901-484A-135	Sequence 135, App
34	154	43.6	182	22	US-09-853-526-133	Sequence 133, App
35	154	43.6	182	23	US-09-901-484A-133	Sequence 133, App
36	153	43.3	176	18	US-09-436-919-5	Sequence 5, Appl1
37	151	42.8	1032	26	US-60-212-413-309	Sequence 309, App
38	151	42.8	1032	26	US-60-229-518-248	Sequence 248, App
39	116	32.9	238	22	US-09-853-526-126	Sequence 126, App
40	116	32.9	238	23	US-09-901-484A-126	Sequence 126, App
41	85	24.1	97	22	US-09-853-526-132	Sequence 132, App
42	85	24.1	97	23	US-09-901-484A-132	Sequence 132, App
43	81	22.9	92	17	US-09-316-123-124	Sequence 124, App
44	81	22.9	92	22	US-09-827-244-124	Sequence 124, App
45	70	19.8	257	1	PCT-US01-08631-31719	Sequence 31719, A

ALIGNMENTS

RESULT 1  
PCT-US01-01431-59  
Sequence 59, Application PC/TUS0101431  
GENERAL INFORMATION:  
APPLICANT: Human Genome Sciences, Inc.  
TITLE OF INVENTION: 17 human secreted proteins  
FILE REFERENCE: PS723PCT  
CURRENT APPLICATION NUMBER: PCT/US01/01431  
CURRENT FILING DATE: 2001-01-17  
PRIOR APPLICATION NUMBER: 60/179,065  
PRIOR FILING DATE: 2000-01-31  
PRIOR APPLICATION NUMBER: 60/180,628  
PRIOR FILING DATE: 2000-02-04  
PRIOR APPLICATION NUMBER: 60/231,968  
NUMBER OF SEQ ID NOS: 97  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 59  
LENGTH: 353  
TYPE: PRT  
ORGANISM: Homo sapiens  
PCT-US01-01431-59

Query Match 100.0%; Score 353; DB 1; Length 353;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 353; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
1 MRLLPSVLLGTAPTYVLAAGVWRLSAFLPARFYQALDDRLCYQOSVLEFFENYTG 60  
|||||

```
Db 1 MRYLLPSVLLGTAPTYVLAMGVWRLLSAFLPARFYQALDRLCYVQSWVLEFFENYTG 60
QY 61 VOILLYDGLPRKNENIYLANHOSTVDVIADILAIROMALGHVRYVLEKGLKWLPLYGC 120
Db 61 VOILLYDGLPRKNENIYLANHOSTVDVIADILAIROMALGHVRYVLEKGLKWLPLYGC 120
QY 121 YFAOHGIIYVRSKAKFNEKEKRNKLOSYVDAGTPMYLVIPPEGTRYNPEQTKVLSASQAF 180
Db 121 YFAOHGIIYVRSKAKFNEKEKRNKLOSYVDAGTPMYLVIPPEGTRYNPEQTKVLSASQAF 180
QY 181 AAOGLAVLKHVLTLPRIKATHVAFDCMKNYLDAYDVTVVYEGKDDGQRRSEPTMTTEFL 240
Db 181 AAOGLAVLKHVLTLPRIKATHVAFDCMKNYLDAYDVTVVYEGKDDGQRRSEPTMTTEFL 240
QY 241 CKECPKTHIHIDRIKDKVDPEQEHMRRLHEREFIKDKMLIEFYESPDPERRRRPFPGKS 300
Db 241 CKECPKTHIHIDRIKDKVDPEQEHMRRLHEREFIKDKMLIEFYESPDPERRRRPFPGKS 300
QY 301 VNSKLSIKTKLPSMLILSGLTAGMLMTDAGRKLVTNTWITGTLGCLMTWITKA 353
Db 301 VNSKLSIKTKLPSMLILSGLTAGMLMTDAGRKLVTNTWITGTLGCLMTWITKA 353
```

```
RESULT 2
PCT-US01-01431-79
; Sequence 79, Application PC/TUS0101431
; GENERAL INFORMATION:
; APPLICANT: Human Genome Sciences, Inc.
; TITLE OF INVENTION: 17 human secreted proteins
; FILE REFERENCE: PS723PCT
; CURRENT APPLICATION NUMBER: PCT/US01/01431
; CURRENT FILING DATE: 2001-01-17
; PRIOR APPLICATION NUMBER: 60/179, 065
; PRIOR FILING DATE: 2000-01-31
; PRIOR APPLICATION NUMBER: 60/180, 628
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: 60/231, 968
; PRIOR FILING DATE: 2000-09-12
; NUMBER OF SEQ ID NOS: 97
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 79
; LENGTH: 353
; TYPE: PRT
; ORGANISM: Homo sapiens
PCT-US01-01431-79
```

```
Query Match 100.0%; Score 353; DB 1; Length 353;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 353; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 MRYLLPSVLLGTAPTYVLAMGVWRLLSAFLPARFYQALDRLCYVQSWVLEFFENYTG 60
Db 1 MRYLLPSVLLGTAPTYVLAMGVWRLLSAFLPARFYQALDRLCYVQSWVLEFFENYTG 60
QY 61 VOILLYDGLPRKNENIYLANHOSTVDVIADILAIROMALGHVRYVLEKGLKWLPLYGC 120
Db 61 VOILLYDGLPRKNENIYLANHOSTVDVIADILAIROMALGHVRYVLEKGLKWLPLYGC 120
QY 121 YFAOHGIIYVRSKAKFNEKEKRNKLOSYVDAGTPMYLVIPPEGTRYNPEQTKVLSASQAF 180
Db 121 YFAOHGIIYVRSKAKFNEKEKRNKLOSYVDAGTPMYLVIPPEGTRYNPEQTKVLSASQAF 180
QY 181 AAOGLAVLKHVLTLPRIKATHVAFDCMKNYLDAYDVTVVYEGKDDGQRRSEPTMTTEFL 240
Db 181 AAOGLAVLKHVLTLPRIKATHVAFDCMKNYLDAYDVTVVYEGKDDGQRRSEPTMTTEFL 240
QY 241 CKECPKTHIHIDRIKDKVDPEQEHMRRLHEREFIKDKMLIEFYESPDPERRRRPFPGKS 300
Db 241 CKECPKTHIHIDRIKDKVDPEQEHMRRLHEREFIKDKMLIEFYESPDPERRRRPFPGKS 300
QY 301 VNSKLSIKTKLPSMLILSGLTAGMLMTDAGRKLVTNTWITGTLGCLMTWITKA 353
Db 301 VNSKLSIKTKLPSMLILSGLTAGMLMTDAGRKLVTNTWITGTLGCLMTWITKA 353
```

```
Db 301 VNSKLSIKTKLPSMLILSGLTAGMLMTDAGRKLVTNTWITGTLGCLMTWITKA 353
RESULT 3
PCT-US01-11988-1474
; Sequence 1474, Application PC/TUS0111988
; GENERAL INFORMATION:
; APPLICANT: Human Genome Sciences, Inc.
; TITLE OF INVENTION: Albumin Fusion Proteins
; FILE REFERENCE: PF546PCT
; CURRENT APPLICATION NUMBER: PCT/US01/11988
; CURRENT FILING DATE: 2001-01-12
; PRIOR APPLICATION NUMBER: 60/229, 358
; PRIOR FILING DATE: 2000-04-12
; PRIOR APPLICATION NUMBER: 60/256, 931
; PRIOR FILING DATE: 2000-12-21
; PRIOR APPLICATION NUMBER: 60/199, 384
; PRIOR FILING DATE: 2000-04-25
; NUMBER OF SEQ ID NOS: 2267
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1474
; LENGTH: 353
; TYPE: PRT
; ORGANISM: Homo sapiens
PCT-US01-11988-1474
```

```
Query Match 100.0%; Score 353; DB 1; Length 353;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 353; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 MRYLLPSVLLGTAPTYVLAMGVWRLLSAFLPARFYQALDRLCYVQSWVLEFFENYTG 60
Db 1 MRYLLPSVLLGTAPTYVLAMGVWRLLSAFLPARFYQALDRLCYVQSWVLEFFENYTG 60
QY 61 VOILLYDGLPRKNENIYLANHOSTVDVIADILAIROMALGHVRYVLEKGLKWLPLYGC 120
Db 61 VOILLYDGLPRKNENIYLANHOSTVDVIADILAIROMALGHVRYVLEKGLKWLPLYGC 120
QY 121 YFAOHGIIYVRSKAKFNEKEKRNKLOSYVDAGTPMYLVIPPEGTRYNPEQTKVLSASQAF 180
Db 121 YFAOHGIIYVRSKAKFNEKEKRNKLOSYVDAGTPMYLVIPPEGTRYNPEQTKVLSASQAF 180
QY 181 AAOGLAVLKHVLTLPRIKATHVAFDCMKNYLDAYDVTVVYEGKDDGQRRSEPTMTTEFL 240
Db 181 AAOGLAVLKHVLTLPRIKATHVAFDCMKNYLDAYDVTVVYEGKDDGQRRSEPTMTTEFL 240
QY 241 CKECPKTHIHIDRIKDKVDPEQEHMRRLHEREFIKDKMLIEFYESPDPERRRRPFPGKS 300
Db 241 CKECPKTHIHIDRIKDKVDPEQEHMRRLHEREFIKDKMLIEFYESPDPERRRRPFPGKS 300
QY 301 VNSKLSIKTKLPSMLILSGLTAGMLMTDAGRKLVTNTWITGTLGCLMTWITKA 353
Db 301 VNSKLSIKTKLPSMLILSGLTAGMLMTDAGRKLVTNTWITGTLGCLMTWITKA 353
```

```
RESULT 4
PCT-US01-11988-1475
; Sequence 1475, Application PC/TUS0111988
; GENERAL INFORMATION:
; APPLICANT: Human Genome Sciences, Inc.
; TITLE OF INVENTION: Albumin Fusion Proteins
; FILE REFERENCE: PF546PCT
; CURRENT APPLICATION NUMBER: PCT/US01/11988
; CURRENT FILING DATE: 2001-01-12
; PRIOR APPLICATION NUMBER: 60/229, 358
; PRIOR FILING DATE: 2000-04-12
; PRIOR APPLICATION NUMBER: 60/256, 931
; PRIOR FILING DATE: 2000-12-21
; PRIOR APPLICATION NUMBER: 60/199, 384
; PRIOR FILING DATE: 2000-04-25
; NUMBER OF SEQ ID NOS: 2267
; SOFTWARE: PatentIn Ver. 2.1
```

SEQ ID NO 1475  
LENGTH: 353  
TYPE: PRT  
ORGANISM: Homo sapiens  
PCT-US01-11988-1475

Query Match 100.0%; Score 353; DB 1; Length 353;  
Best local Similarity 100.0%; Pred. No. 0;  
Matches 353; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MRYLPSVLLGTAPTYVILAMGWRLISAFLPARFYQALDDRLCYVQSMVLEFFENYTG 60  
DB 1 MRYLPSVLLGTAPTYVILAMGWRLISAFLPARFYQALDDRLCYVQSMVLEFFENYTG 60  
QY 61 VOILLYGDLPKKNENIYYLANHOSYDWMIVADILAIQNALGHVRYLAKGKMLPLYGC 120  
DB 61 VOILLYGDLPKKNENIYYLANHOSYDWMIVADILAIQNALGHVRYLAKGKMLPLYGC 120  
QY 121 YFAOHGIIYKRSKAKFNEKEMRNKLOSVDAGTPMYLVIPEEGTRVNPEDTKVLSAQAF 180  
DB 121 YFAOHGIIYKRSKAKFNEKEMRNKLOSVDAGTPMYLVIPEEGTRVNPEDTKVLSAQAF 180  
QY 181 AARGLAIVLKHVLTLPRIKATHVAFDCMKNYLDAIYDVTYVEGKDDGGORRESPTMTFEL 240  
DB 181 AARGLAIVLKHVLTLPRIKATHVAFDCMKNYLDAIYDVTYVEGKDDGGORRESPTMTFEL 240  
QY 241 CKECPRKIHIDRIDKDDVEEEOEHMRMWLHERFEIKDKMLIEFYESPDERRRKRPFGKS 300  
DB 241 CKECPRKIHIDRIDKDDVEEEOEHMRMWLHERFEIKDKMLIEFYESPDERRRKRPFGKS 300  
QY 301 VNSKLSIKKTLPSMLILSGLTAGMLMTDAGRKLIVNTWYIGTLGCLMWTIKA 353  
DB 301 VNSKLSIKKTLPSMLILSGLTAGMLMTDAGRKLIVNTWYIGTLGCLMWTIKA 353

RESULT 5  
US-09-833-245-1474  
; Sequence 1474, Application US/09833245  
; GENERAL INFORMATION:  
; APPLICANT: Human Genome Sciences, Inc.  
; TITLE OF INVENTION: Albumin Fusion Proteins  
; FILE REFERENCE: PF546PCT  
; CURRENT APPLICATION NUMBER: US/09/833,245  
; PRIOR FILING DATE: 2001-04-12  
; PRIOR APPLICATION NUMBER: 60/229, 358  
; PRIOR FILING DATE: 2000-04-12  
; PRIOR APPLICATION NUMBER: 60/256, 931  
; PRIOR FILING DATE: 2000-12-21  
; PRIOR APPLICATION NUMBER: 60/199, 384  
; PRIOR FILING DATE: 2000-04-25  
; NUMBER OF SEQ ID NOS: 2267  
; SOFTWARE: Patentln Ver. 2.1  
; SEQ ID NO 1474  
; LENGTH: 353  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-833-245-1474

Query Match 100.0%; Score 353; DB 22; Length 353;  
Best local Similarity 100.0%; Pred. No. 0;  
Matches 353; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 MRYLPSVLLGTAPTYVILAMGWRLISAFLPARFYQALDDRLCYVQSMVLEFFENYTG 60  
DB 1 MRYLPSVLLGTAPTYVILAMGWRLISAFLPARFYQALDDRLCYVQSMVLEFFENYTG 60  
QY 61 VOILLYGDLPKKNENIYYLANHOSYDWMIVADILAIQNALGHVRYLAKGKMLPLYGC 120  
DB 61 VOILLYGDLPKKNENIYYLANHOSYDWMIVADILAIQNALGHVRYLAKGKMLPLYGC 120  
QY 121 YFAOHGIIYKRSKAKFNEKEMRNKLOSVDAGTPMYLVIPEEGTRVNPEDTKVLSAQAF 180

DB 121 YFAOHGIIYKRSKAKFNEKEMRNKLOSVDAGTPMYLVIPEEGTRVNPEDTKVLSAQAF 180  
QY 181 AARGLAIVLKHVLTLPRIKATHVAFDCMKNYLDAIYDVTYVEGKDDGGORRESPTMTFEL 240  
DB 181 AARGLAIVLKHVLTLPRIKATHVAFDCMKNYLDAIYDVTYVEGKDDGGORRESPTMTFEL 240  
QY 241 CKECPRKIHIDRIDKDDVEEEOEHMRMWLHERFEIKDKMLIEFYESPDERRRKRPFGKS 300  
DB 241 CKECPRKIHIDRIDKDDVEEEOEHMRMWLHERFEIKDKMLIEFYESPDERRRKRPFGKS 300  
QY 301 VNSKLSIKKTLPSMLILSGLTAGMLMTDAGRKLIVNTWYIGTLGCLMWTIKA 353  
DB 301 VNSKLSIKKTLPSMLILSGLTAGMLMTDAGRKLIVNTWYIGTLGCLMWTIKA 353

RESULT 6  
US-09-833-245-1475  
; Sequence 1475, Application US/09833245  
; GENERAL INFORMATION:  
; APPLICANT: Human Genome Sciences, Inc.  
; TITLE OF INVENTION: Albumin Fusion Proteins  
; FILE REFERENCE: PF546PCT  
; CURRENT APPLICATION NUMBER: US/09/833,245  
; PRIOR FILING DATE: 2001-04-12  
; PRIOR APPLICATION NUMBER: 60/229, 358  
; PRIOR FILING DATE: 2000-04-12  
; PRIOR APPLICATION NUMBER: 60/256, 931  
; PRIOR FILING DATE: 2000-12-21  
; PRIOR APPLICATION NUMBER: 60/199, 384  
; PRIOR FILING DATE: 2000-04-25  
; NUMBER OF SEQ ID NOS: 2267  
; SOFTWARE: Patentln Ver. 2.1  
; SEQ ID NO 1475  
; LENGTH: 353  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-833-245-1475

Query Match 100.0%; Score 353; DB 22; Length 353;  
Best local Similarity 100.0%; Pred. No. 0;  
Matches 353; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 MRYLPSVLLGTAPTYVILAMGWRLISAFLPARFYQALDDRLCYVQSMVLEFFENYTG 60  
DB 1 MRYLPSVLLGTAPTYVILAMGWRLISAFLPARFYQALDDRLCYVQSMVLEFFENYTG 60  
QY 61 VOILLYGDLPKKNENIYYLANHOSYDWMIVADILAIQNALGHVRYLAKGKMLPLYGC 120  
DB 61 VOILLYGDLPKKNENIYYLANHOSYDWMIVADILAIQNALGHVRYLAKGKMLPLYGC 120  
QY 121 YFAOHGIIYKRSKAKFNEKEMRNKLOSVDAGTPMYLVIPEEGTRVNPEDTKVLSAQAF 180  
DB 121 YFAOHGIIYKRSKAKFNEKEMRNKLOSVDAGTPMYLVIPEEGTRVNPEDTKVLSAQAF 180  
QY 181 AARGLAIVLKHVLTLPRIKATHVAFDCMKNYLDAIYDVTYVEGKDDGGORRESPTMTFEL 240  
DB 181 AARGLAIVLKHVLTLPRIKATHVAFDCMKNYLDAIYDVTYVEGKDDGGORRESPTMTFEL 240  
QY 241 CKECPRKIHIDRIDKDDVEEEOEHMRMWLHERFEIKDKMLIEFYESPDERRRKRPFGKS 300  
DB 241 CKECPRKIHIDRIDKDDVEEEOEHMRMWLHERFEIKDKMLIEFYESPDERRRKRPFGKS 300  
QY 301 VNSKLSIKKTLPSMLILSGLTAGMLMTDAGRKLIVNTWYIGTLGCLMWTIKA 353  
DB 301 VNSKLSIKKTLPSMLILSGLTAGMLMTDAGRKLIVNTWYIGTLGCLMWTIKA 353  
RESULT 7  
US-09-853-526-4  
; Sequence 4, Application US/09853526  
; GENERAL INFORMATION:

APPLICANT: Cohen, Daniel  
APPLICANT: Blumenfeld, Marta  
APPLICANT: Ilya, Chumakov  
APPLICANT: Bougueleret, Lydie  
TITLE OF INVENTION: PROSTATE CANCER GENE  
FILE REFERENCE: GENSET.18CPICP  
CURRENT APPLICATION NUMBER: US/09/853,526  
PRIORITY FILING DATE: 2001-05-11  
PRIORITY APPLICATION NUMBER: 09/338,907  
PRIORITY FILING DATE: 1999-06-23  
PRIORITY APPLICATION NUMBER: 08/996,306  
PRIORITY FILING DATE: 1997-12-22  
PRIORITY APPLICATION NUMBER: 60/099,658  
PRIORITY FILING DATE: 1998-09-09  
PRIORITY APPLICATION NUMBER: 09/218,207  
PRIORITY FILING DATE: 1998-12-22  
NUMBER OF SEQ ID NOS: 578  
SOFTWARE: Patent.pm  
SEQ ID NO 4  
LENGTH: 353  
TYPE: PRT  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: HELIX  
LOCATION: 1..33  
OTHER INFORMATION: Rao and Argos identification method, potential helix  
NAME/KEY: HELIX  
LOCATION: 4..20  
OTHER INFORMATION: Klein, Kanehisa and Delisi identification method, potential  
NAME/KEY: HELIX  
LOCATION: 4..24  
OTHER INFORMATION: Eisenberg, Schwarz, Komarony, Wall identification method, F  
NAME/KEY: MYRISTATE  
LOCATION: 12..16  
OTHER INFORMATION: Prosite match  
NAME/KEY: HELIX  
LOCATION: 50..70  
OTHER INFORMATION: Eisenberg, Schwarz, Komarony, Wall identification method, F  
NAME/KEY: CARBOHYD  
LOCATION: 57..59  
OTHER INFORMATION: Prosite match  
NAME/KEY: HELIX  
LOCATION: 76..96  
OTHER INFORMATION: Eisenberg, Schwarz, Komarony, Wall identification method, F  
NAME/KEY: PHOSPHORYLATION  
LOCATION: 84  
OTHER INFORMATION: potential Tyrosine kinase site, Prosite match  
NAME/KEY: SITE  
LOCATION: 94..115  
OTHER INFORMATION: potential caseine kinase II site, Prosite match  
NAME/KEY: MYRISTATE  
LOCATION: 119..123  
OTHER INFORMATION: potential Leucine zipper site, Prosite match  
NAME/KEY: PHOSPHORYLATION  
LOCATION: 133  
OTHER INFORMATION: potential site, Prosite match  
NAME/KEY: PHOSPHORYLATION  
LOCATION: 147  
OTHER INFORMATION: potential protein kinase C, Prosite match  
NAME/KEY: PHOSPHORYLATION  
LOCATION: 194  
OTHER INFORMATION: potential caseine kinase II site, Prosite match  
NAME/KEY: PHOSPHORYLATION  
LOCATION: 215  
OTHER INFORMATION: potential protein kinase C, Prosite match  
NAME/KEY: SULFATATION  
LOCATION: 221  
OTHER INFORMATION: potential Tyrosine kinase site, Prosite match

OTHER INFORMATION: Prosite match  
NAME/KEY: PHOSPHORYLATION  
LOCATION: 233  
OTHER INFORMATION: potential CAMP and cGMP dependant protein kinase site,  
OTHER INFORMATION: match  
NAME/KEY: PHOSPHORYLATION  
LOCATION: 235  
OTHER INFORMATION: potential caseine kinase II site, Prosite match  
NAME/KEY: PHOSPHORYLATION  
LOCATION: 306  
OTHER INFORMATION: potential protein kinase C, Prosite match  
NAME/KEY: HELIX  
LOCATION: 310..330  
OTHER INFORMATION: Eisenberg, Schwarz, Komarony, Wall identification method,  
OTHER INFORMATION: helix  
NAME/KEY: MYRISTATE  
LOCATION: 319..323  
OTHER INFORMATION: Prosite match  
NAME/KEY: MYRISTATE  
LOCATION: 323..327  
OTHER INFORMATION: Prosite match  
NAME/KEY: AMIDATION  
LOCATION: 329  
OTHER INFORMATION: Prosite match  
NAME/KEY: HELIX  
LOCATION: 333..353  
OTHER INFORMATION: Eisenberg, Schwarz, Komarony, Wall identification method,  
OTHER INFORMATION: helix  
NAME/KEY: MYRISTATE  
LOCATION: 341..345  
OTHER INFORMATION: Prosite match  
NAME/KEY: PHOSPHORYLATION  
LOCATION: 350  
OTHER INFORMATION: potential protein kinase C, Prosite match  
US-09-853-526-4

Query Match 100.0%; Score 353; DB 22; Length 353;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 353; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MRYLLPSVVLGTPATYVYVLMGWRLSAPLPARYOALDDRLCYQVOSVWLFEEFYTG 60  
Db 1 MRYLLPSVVLGTPATYVYVLMGWRLSAPLPARYOALDDRLCYQVOSVWLFEEFYTG 60  
QY 61 VOILLYDLPKKNENIYYLANHOSYVDMIVADILAIQNALGHVRYVLYKGLKMLPLYGC 120  
Db 61 VOILLYDLPKKNENIYYLANHOSYVDMIVADILAIQNALGHVRYVLYKGLKMLPLYGC 120  
QY 121 YFAOHGSIYVRSKAKFNEKEMRNKLOSYVDAGTPMYLVIFPEGTRVNPEDTKVLSAQAF 180  
Db 121 YFAOHGSIYVRSKAKFNEKEMRNKLOSYVDAGTPMYLVIFPEGTRVNPEDTKVLSAQAF 180  
QY 181 AAOEGGLAVLKHVLTTPRKATNVHAFDCKKNYLDVIVDTVYVEKDKDGGORRESPTMEFL 240  
Db 181 AAOEGGLAVLKHVLTTPRKATNVHAFDCKKNYLDVIVDTVYVEKDKDGGORRESPTMEFL 240  
QY 241 CKECPKIHIDRIDKQVDEDEHMRWLHEREFEIKDKMLIEFYESPDERRRKRPFGKS 300  
Db 241 CKECPKIHIDRIDKQVDEDEHMRWLHEREFEIKDKMLIEFYESPDERRRKRPFGKS 300  
QY 301 VNSKLSIKITLPSMLIISGLTAGMLMTDAGRKLIVNTWITGTLGCIWPTIKA 353  
Db 301 VNSKLSIKITLPSMLIISGLTAGMLMTDAGRKLIVNTWITGTLGCIWPTIKA 353

RESULT 8  
US-09-901-484A-4  
Sequence 4, Application US/09901484A  
GENERAL INFORMATION:  
APPLICANT: Cohen, Daniel  
APPLICANT: Blumenfeld, Marta  
APPLICANT: Chumakov, Ilya



APPLICANT: Bougueleret, Lydie  
TITLE OF INVENTION: Prostate Cancer Gene  
FILE REFERENCE: GEN-T11XC3D2  
CURRENT APPLICATION NUMBER: US/09/901,484A  
CURRENT FILING DATE: 2001-07-09  
PRIOR APPLICATION NUMBER: US 08/996,306  
PRIOR FILING DATE: 1997-12-22  
PRIOR APPLICATION NUMBER: US 60/099,658  
PRIOR FILING DATE: 1998-09-09  
PRIOR APPLICATION NUMBER: US 09/218,207  
PRIOR FILING DATE: 1998-12-22  
PRIOR APPLICATION NUMBER: US 09/338,907  
PRIOR FILING DATE: 1999-06-23  
PRIOR APPLICATION NUMBER: US 09/853,526  
PRIOR FILING DATE: 2001-05-11  
NUMBER OF SEQ ID NOS: 578  
SOFTWARE: PatentIn version 3.1  
SEQ ID NO 4  
LENGTH: 353  
TYPE: PRT  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: HELIX  
LOCATION: (1)..(33)  
OTHER INFORMATION: Rao and Argos identification method, potential helix  
NAME/KEY: HELIX  
LOCATION: (4)..(20)  
OTHER INFORMATION: Klein, Kanehisa and Delisi identification method, potential helix  
NAME/KEY: HELIX  
LOCATION: (4)..(24)  
OTHER INFORMATION: Eisenberg, Schwarz, Komarony, Wall identification method,  
OTHER INFORMATION: potential helix  
NAME/KEY: LIPID  
LOCATION: (12)..(16)  
OTHER INFORMATION: MYRISTATE, Prosite match  
NAME/KEY: HELIX  
LOCATION: (50)..(70)  
OTHER INFORMATION: Eisenberg, Schwarz, Komarony, Wall identification method,  
OTHER INFORMATION: potential helix  
NAME/KEY: CARBOHYD  
LOCATION: (57)..(59)  
OTHER INFORMATION: Prosite match  
NAME/KEY: HELIX  
LOCATION: (76)..(96)  
OTHER INFORMATION: Eisenberg, Schwarz, Komarony, Wall identification method,  
OTHER INFORMATION: potential helix  
NAME/KEY: MOD\_RES  
LOCATION: (78)..(78)  
OTHER INFORMATION: PHOSPHORYLATION, potential Tyrosine kinase site, Prosite match  
NAME/KEY: MOD\_RES  
LOCATION: (84)..(84)  
OTHER INFORMATION: PHOSPHORYLATION, potential caseine kinase II site, Prosite match  
NAME/KEY: SITE  
LOCATION: (94)..(115)  
OTHER INFORMATION: Potential leucine zipper site, Prosite match  
NAME/KEY: LIPID  
LOCATION: (119)..(123)  
OTHER INFORMATION: MYRISTATE, Prosite match  
NAME/KEY: MOD\_RES  
LOCATION: (133)..(133)  
OTHER INFORMATION: PHOSPHORYLATION, potential protein kinase C, Prosite match  
NAME/KEY: MOD\_RES  
LOCATION: (147)..(147)  
OTHER INFORMATION: PHOSPHORYLATION, potential caseine kinase II, Prosite match  
NAME/KEY: MOD\_RES  
LOCATION: (194)..(194)  
OTHER INFORMATION: PHOSPHORYLATION, potential protein kinase C, Prosite match  
NAME/KEY: MOD\_RES  
LOCATION: (215)..(215)  
OTHER INFORMATION: PHOSPHORYLATION, potential Tyrosine kinase site, Prosite match  
NAME/KEY: MOD\_RES  
LOCATION: (221)..(221)  
OTHER INFORMATION: SULFATATION, Prosite match

NAME/KEY: MOD\_RES  
LOCATION: (233)..(233)  
OTHER INFORMATION: PHOSPHORYLATION, potential CAMP and cGMP dependant protein klr  
OTHER INFORMATION: site, Prosite match  
NAME/KEY: MOD\_RES  
LOCATION: (235)..(235)  
OTHER INFORMATION: PHOSPHORYLATION, potential caseine kinase II site, Prosite mat  
NAME/KEY: MOD\_RES  
LOCATION: (306)..(306)  
OTHER INFORMATION: PHOSPHORYLATION, potential protein kinase C, Prosite match  
NAME/KEY: HELIX  
LOCATION: (310)..(330)  
OTHER INFORMATION: Eisenberg, Schwarz, Komarony, Wall identification method,  
OTHER INFORMATION: potential helix  
NAME/KEY: LIPID  
LOCATION: (319)..(323)  
OTHER INFORMATION: MYRISTATE, Prosite match  
NAME/KEY: LIPID  
LOCATION: (323)..(327)  
OTHER INFORMATION: MYRISTATE, Prosite match  
NAME/KEY: MOD\_RES  
LOCATION: (329)..(329)  
OTHER INFORMATION: AMIDATION, Prosite match  
NAME/KEY: HELIX  
LOCATION: (333)..(353)  
OTHER INFORMATION: Eisenberg, Schwarz, Komarony, Wall identification method,  
OTHER INFORMATION: potential helix  
NAME/KEY: LIPID  
LOCATION: (341)..(345)  
OTHER INFORMATION: MYRISTATE, Prosite match  
NAME/KEY: MOD\_RES  
LOCATION: (350)..(350)  
OTHER INFORMATION: PHOSPHORYLATION, potential protein kinase C, Prosite match  
US-09-901-484A-4

Query Match 100.0%; Score 353; DB 23; Length 353;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 353; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 MRYLLPSVLLGTAPTYVLMGWRLLSAFLEPAFYQALDDRLCYVQSWLEFFENYTG 60  
DB 1 MRYLLPSVLLGTAPTYVLMGWRLLSAFLEPAFYQALDDRLCYVQSWLEFFENYTG 60  
QY 61 VOILLYDDLPKKNENIYYLANHOSVDVYADILATQNLGHHVRYLKGKWLPLYGC 120  
DB 61 VOILLYDDLPKKNENIYYLANHOSVDVYADILATQNLGHHVRYLKGKWLPLYGC 120  
QY 121 YFAOHGSIYVRSKAFNEKEMRNKLOSYVDAGTPMYLVIPPEGTRVYPEOTKYLASQAF 180  
DB 121 YFAOHGSIYVRSKAFNEKEMRNKLOSYVDAGTPMYLVIPPEGTRVYPEOTKYLASQAF 180  
QY 181 AAOGLAVLKHVLTPIKATHVAFDCMKNYLDALYDVTVVYEGKDGQRRSEPTWTEFL 240  
DB 181 AAOGLAVLKHVLTPIKATHVAFDCMKNYLDALYDVTVVYEGKDGQRRSEPTWTEFL 240  
QY 241 CKCEPKTHIHIDRDKKDVEEOEHMRMYLHEREPEIKDKMLIEYESPDDERRRRFPFGKS 300  
DB 241 CKCEPKTHIHIDRDKKDVEEOEHMRMYLHEREPEIKDKMLIEYESPDDERRRRFPFGKS 300  
QY 301 VNSKLSIKTLPMSMLISGLTAGMLMTDAGRKYLVNTWYIGTLLGCLMTYIKA 353  
DB 301 VNSKLSIKTLPMSMLISGLTAGMLMTDAGRKYLVNTWYIGTLLGCLMTYIKA 353

RESULT 9  
US-09-915-582-59  
Sequence 59, Application US/09915582  
GENERAL INFORMATION:  
APPLICANT: Rosen et al.  
TITLE OF INVENTION: 17 Human Secreted Proteins  
FILE REFERENCE: PS723PI  
CURRENT APPLICATION NUMBER: US/09/915,582

;; CURRENT FILING DATE: 2001-07-27  
;; PRIOR APPLICATION NUMBER: PCT/US01/01431  
;; PRIOR FILING DATE: 2001-01-17  
;; PRIOR APPLICATION NUMBER: 60/179,065  
;; PRIOR FILING DATE: 2000-01-31  
;; PRIOR APPLICATION NUMBER: 60/180,628  
;; PRIOR FILING DATE: 2000-02-04  
;; PRIOR APPLICATION NUMBER: 60/231,968  
;; PRIOR FILING DATE: 2000-09-12  
;; NUMBER OF SEQ ID NOS: 97  
;; SOFTWARE: PatentIn Ver. 2.0  
;; SEQ ID NO 59  
;; TYPE: PRT  
;; ORGANISM: Homo sapiens  
US-09-915-582-59

Query Match 100.0%; Score 353; DB 23; Length 353;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 353; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 MRYLLPSVLLGTAPTYVLAWGVRLLSAFLPARFYQALDDRLCYVQSWLFFFEENTG 60  
DB 1 MRYLLPSVLLGTAPTYVLAWGVRLLSAFLPARFYQALDDRLCYVQSWLFFFEENTG 60  
QY 61 VQILLYDLPKKNENIYLANHSTVDMIVADILAIQNALGHRYVLEKGLKMLPLYGC 120  
DB 61 VQILLYDLPKKNENIYLANHSTVDMIVADILAIQNALGHRYVLEKGLKMLPLYGC 120  
QY 121 YFAOHGCIYVRSKAFNEKEKRNKLOSVDAGTMYLVIPEEGTRYNPEOTKVLASQAF 180  
DB 121 YFAOHGCIYVRSKAFNEKEKRNKLOSVDAGTMYLVIPEEGTRYNPEOTKVLASQAF 180  
QY 181 AAORGGLAVLKVLTPRIKATHVAFDCMKNYDAIYDVTVYVEGKDDGQRRSEPTMTPEFL 240  
DB 181 AAORGGLAVLKVLTPRIKATHVAFDCMKNYDAIYDVTVYVEGKDDGQRRSEPTMTPEFL 240  
QY 241 CKECPKIHIDRIDKKDVPEQEHMRRLHEREIKDKMLIEFYESPDERRRKRPFGKS 300  
DB 241 CKECPKIHIDRIDKKDVPEQEHMRRLHEREIKDKMLIEFYESPDERRRKRPFGKS 300  
QY 301 VNSKLSIKTKLPMLISGLTAGMLMTDAGRKLYVNTWYIGTLGCLMWTIKA 353  
DB 301 VNSKLSIKTKLPMLISGLTAGMLMTDAGRKLYVNTWYIGTLGCLMWTIKA 353

RESULT 10  
US-09-915-582-79  
;; Sequence 79, Application US/09915582  
;; GENERAL INFORMATION:  
;; APPLICANT: Rosen et al.  
;; TITLE OF INVENTION: 17 Human Secreted Proteins  
;; FILE REFERENCE: P5723P1  
;; CURRENT APPLICATION NUMBER: US/09/915,582  
;; PRIOR FILING DATE: 2001-07-27  
;; PRIOR APPLICATION NUMBER: PCT/US01/01431  
;; PRIOR FILING DATE: 2001-01-17  
;; PRIOR APPLICATION NUMBER: 60/179,065  
;; PRIOR FILING DATE: 2000-01-31  
;; PRIOR APPLICATION NUMBER: 60/180,628  
;; PRIOR FILING DATE: 2000-02-04  
;; PRIOR APPLICATION NUMBER: 60/231,968  
;; PRIOR FILING DATE: 2000-09-12  
;; NUMBER OF SEQ ID NOS: 97  
;; SOFTWARE: PatentIn Ver. 2.0  
;; SEQ ID NO 79  
;; LENGTH: 353  
;; TYPE: PRT  
;; ORGANISM: Homo sapiens  
US-09-915-582-79

Query Match 100.0%; Score 353; DB 23; Length 353;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 353; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 MRYLLPSVLLGTAPTYVLAWGVRLLSAFLPARFYQALDDRLCYVQSWLFFFEENTG 60  
DB 1 MRYLLPSVLLGTAPTYVLAWGVRLLSAFLPARFYQALDDRLCYVQSWLFFFEENTG 60  
QY 61 VQILLYDLPKKNENIYLANHSTVDMIVADILAIQNALGHRYVLEKGLKMLPLYGC 120  
DB 61 VQILLYDLPKKNENIYLANHSTVDMIVADILAIQNALGHRYVLEKGLKMLPLYGC 120  
QY 121 YFAOHGCIYVRSKAFNEKEKRNKLOSVDAGTMYLVIPEEGTRYNPEOTKVLASQAF 180  
DB 121 YFAOHGCIYVRSKAFNEKEKRNKLOSVDAGTMYLVIPEEGTRYNPEOTKVLASQAF 180  
QY 181 AAORGGLAVLKVLTPRIKATHVAFDCMKNYDAIYDVTVYVEGKDDGQRRSEPTMTPEFL 240  
DB 181 AAORGGLAVLKVLTPRIKATHVAFDCMKNYDAIYDVTVYVEGKDDGQRRSEPTMTPEFL 240  
QY 241 CKECPKIHIDRIDKKDVPEQEHMRRLHEREIKDKMLIEFYESPDERRRKRPFGKS 300  
DB 241 CKECPKIHIDRIDKKDVPEQEHMRRLHEREIKDKMLIEFYESPDERRRKRPFGKS 300  
QY 301 VNSKLSIKTKLPMLISGLTAGMLMTDAGRKLYVNTWYIGTLGCLMWTIKA 353  
DB 301 VNSKLSIKTKLPMLISGLTAGMLMTDAGRKLYVNTWYIGTLGCLMWTIKA 353

RESULT 11  
US-60-099-658-4  
;; Sequence 4, Application US/60099658  
;; GENERAL INFORMATION:  
;; APPLICANT: Cohen, Daniel  
;; APPLICANT: Chumakov, Ilya  
;; APPLICANT: Blumenfeld, Marta  
;; APPLICANT: Bougueterec, Lydie  
;; TITLE OF INVENTION: Prostate cancer gene  
;; NUMBER OF SEQUENCES: 99  
;; CORRESPONDENCE ADDRESS:  
;; ADDRESSEE: Knobbe, Martens, Olson & Bear  
;; STREET: 501 West Broadway  
;; CITY: San Diego  
;; STATE: California  
;; COUNTRY: USA  
;; ZIP: 92101-3505  
;; COMPUTER READABLE FORM:  
;; MEDIUM TYPE: Floppy disk  
;; COMPUTER: IBM PC compatible  
;; OPERATING SYSTEM: Win95  
;; SOFTWARE: Word  
;; CURRENT APPLICATION DATA:  
;; APPLICATION NUMBER: US/60/099,658  
;; FILING DATE:  
;; CLASSIFICATION:  
;; ATTORNEY/AGENT INFORMATION:  
;; NAME: Israelson, Ned A.  
;; REGISTRATION NUMBER: 29,655  
;; REFERENCE/DOCKET NUMBER: GENSET.018APR  
;; TELEPHONE: (619) 235-8550  
;; TELEFAX: (619) 235-0176  
;; INFORMATION FOR SEQ ID NO: 4:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 353 amino acids  
;; TYPE: AMINO ACID  
;; STRANDEDNESS: SINGLE  
;; TOPOLOGY: LINEAR  
;; MOLECULE TYPE: PROTEIN  
;; ORIGINAL SOURCE:  
;; ORGANISM: Homo sapiens  
;; FEATURE:  
;; NAME/KEY: potential Transmembrane helix

LOCATION: 1..33	IDENTIFICATION METHOD: Rao and Argos method
FEATURE:	
NAME/KEY: potential Transmembrane helix	
LOCATION: 4..20	
IDENTIFICATION METHOD: Klein, Kanehisa and Delisi method	
FEATURE:	
NAME/KEY: potential Transmembrane helix	
LOCATION: 4..24	
IDENTIFICATION METHOD: Eisenberg,Schwarz,Komarony	
IDENTIFICATION METHOD: and Wall method	
FEATURE:	
NAME/KEY: potential N-myristoylation site	
LOCATION: 12	
IDENTIFICATION METHOD: prosite match	
FEATURE:	
NAME/KEY: potential Transmembrane helix	
LOCATION: 50..70	
IDENTIFICATION METHOD: Eisenberg,Schwarz,Komarony and Wall method	
FEATURE:	
NAME/KEY: potential N-glycosylation site	
LOCATION: 57	
IDENTIFICATION METHOD: prosite match	
FEATURE:	
NAME/KEY: potential Transmembrane helix	
LOCATION: 76..96	
IDENTIFICATION METHOD: Eisenberg,Schwarz,Komarony and Wall method	
FEATURE:	
NAME/KEY: potential Tyrosine kinase phosphorylation site	
LOCATION: 78	
IDENTIFICATION METHOD: prosite match	
FEATURE:	
NAME/KEY: potential Casein kinase II phosphorylation site	
LOCATION: 84	
IDENTIFICATION METHOD: prosite match	
FEATURE:	
NAME/KEY: potential Leucine zipper pattern	
LOCATION: 94..115	
IDENTIFICATION METHOD: prosite match	
FEATURE:	
NAME/KEY: potential N-myristoylation site	
LOCATION: 119	
IDENTIFICATION METHOD: prosite match	
FEATURE:	
NAME/KEY: potential Protein kinase C phosphorylation site	
LOCATION: 133	
IDENTIFICATION METHOD: prosite match	
FEATURE:	
NAME/KEY: potential Casein kinase II phosphorylation site	
LOCATION: 147	
IDENTIFICATION METHOD: prosite match	
FEATURE:	
NAME/KEY: potential protein kinase C phosphorylation site	
LOCATION: 194	
IDENTIFICATION METHOD: prosite match	
FEATURE:	
NAME/KEY: potential Tyrosine kinase phosphorylation site	
LOCATION: 215	
IDENTIFICATION METHOD: prosite match	
FEATURE:	
NAME/KEY: potential Tyrosine sulfatation site	
LOCATION: 221	
IDENTIFICATION METHOD: prosite match	
FEATURE:	
NAME/KEY: potential cAMP- and cGMP-dependent protein kinase phosphorylation site	
LOCATION: 233	
IDENTIFICATION METHOD: prosite match	
FEATURE:	
NAME/KEY: potential Casein kinase II phosphorylation site	
LOCATION: 235	
IDENTIFICATION METHOD: prosite match	
FEATURE:	
NAME/KEY: potential Protein kinase C phosphorylation site	

```

1 LOCATION: 306
2 IDENTIFICATION METHOD: prosite match
3 FEATURE:
4 NAME/KEY: potential Transmembrane helix
5 LOCATION: 310..330
6 IDENTIFICATION METHOD: Eisenberg,Schwarz,Komarony and Wall method
7 FEATURE:
8 NAME/KEY: potential N-myristoylation site
9 LOCATION: 319
10 IDENTIFICATION METHOD: prosite match
11 FEATURE:
12 NAME/KEY: potential N-myristoylation site
13 LOCATION: 323
14 IDENTIFICATION METHOD: prosite match
15 FEATURE:
16 NAME/KEY: potential Amidation site
17 LOCATION: 329
18 IDENTIFICATION METHOD: prosite match
19 FEATURE:
20 NAME/KEY: potential Transmembrane helix
21 LOCATION: 333..353
22 IDENTIFICATION METHOD: Eisenberg,Schwarz,Komarony and Wall method
23 FEATURE:
24 NAME/KEY: potential N-myristoylation site
25 LOCATION: 341
26 IDENTIFICATION METHOD: prosite match
27 FEATURE:
28 NAME/KEY: potential Protein kinase C phosphorylation site
29 LOCATION: 350
30 IDENTIFICATION METHOD: prosite match
31 IS-60-099-658-4

```

	Query Match	Similarity	100.0%	Score 353	DB 26	Length 353
	Best Local	Similarity	100.0%	Pred. No. 0		
	Matches 353	Conservative	0	Mismatches	0	Indels 0
						Gaps
QY	1	MRLLPSVLLGTAPTYVLANGVRRLLSAFLPARFYQALDDRLCYVQSNWLPFFENYTG	60			
Db	1	MRLLPSVLLGTAPTYVLANGVRRLLSAFLPARFYQALDDRLCYVQSNWLPFFENYTG	60			
QY	61	VQILLGDLPAKNENIYYLANHOSVDMIVADILAIKONALGHRVYVLKEGLKMLPLYGC	120			
Db	61	VQILLGDLPAKNENIYYLANHOSVDMIVADILAIKONALGHRVYVLKEGLKMLPLYGC	120			
QY	121	YFAOHGGIYVRSKAKFENEKEMRNKLSQSYVDAGTDMYLVIEPEGTRVNEQTKVLASQAF	180			
Db	121	YFAOHGGIYVRSKAKFENEKEMRNKLSQSYVDAGTDMYLVIEPEGTRVNEQTKVLASQAF	180			
QY	181	AAQGLAHLKHLVLPRIKATFVARDCKNLYDAIYDVTYVEGDDGGQRESPTMTEFL	240			
Db	181	AAQGLAHLKHLVLPRIKATFVARDCKNLYDAIYDVTYVEGDDGGQRESPTMTEFL	240			
QY	241	CKECPKIHIIHDIRDKKDVPDEQEHMRMRLLHEPFEIKDKMLIEFYESDPDERKKRPFGKS	300			
Db	241	CKECPKIHIIHDIRDKKDVPDEQEHMRMRLLHEPFEIKDKMLIEFYESDPDERKKRPFGKS	300			
QY	301	VNSKLSIKTKLPSMLISGLTAGMLMTDAGRKLVVNTWIIYGTLLGCLAWTIKA	353			
Db	301	VNSKLSIKTKLPSMLISGLTAGMLMTDAGRKLVVNTWIIYGTLLGCLAWTIKA	353			

RESULT 12  
 US-09-436-919-1  
 Sequence 1, Application US/09436919A  
 GENERAL INFORMATION:  
 APPLICANT: Leung, David W  
 TITLE OF INVENTION: Human Lysophosphatidic Acid Acyltransferase-epsilon  
 FILE REFERENCE: 1801B  
 CURRENT APPLICATION NUMBER: US/09/436, 919A  
 CURRENT FILING DATE: 1999-11-09  
 NUMBER OF SEQ ID NOS: 5  
 SOFTWARE: Patentin Ver. 2.0

```
; SEQ ID NO 1
; LENGTH: 364
; TYPE: PRF
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: hLRPATeps110
US-09-436-919-1
```

```
Query Match          100.0%; Score 353; DB 18; Length 364;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 353; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY 1 MRYLLPSVLLGTAPTYVLAMGWRLLSAFLPARFYQALDRLCYVQSWLFFFEYNTG 60
    |||||||
DB 12 MRYLLPSVLLGTAPTYVLAMGWRLLSAFLPARFYQALDRLCYVQSWLFFFEYNTG 71
    |||||||
OY 61 VOILLYGDLPKKNENIYYLANHOSTVDMIVADILAIQNALGHRVYLYKGLKMLPLYGC 120
    |||||||
DB 72 VOILLYGDLPKKNENIYYLANHOSTVDMIVADILAIQNALGHRVYLYKGLKMLPLYGC 131
    |||||||
OY 121 YFAOHGIIYVRSKAKFENEKEMRNKLOSVDAGTPMYLVIFPEGTRYNPEOTKVLASQAQF 180
    |||||||
DB 132 YFAOHGIIYVRSKAKFENEKEMRNKLOSVDAGTPMYLVIFPEGTRYNPEOTKVLASQAQF 191
    |||||||
OY 181 AAOGLAVLKHVLTLPRIKATHVAFDCMKNYLDALYDVTYVEGKDGGQRRSEPTMTTEFL 240
    |||||||
DB 192 AAOGLAVLKHVLTLPRIKATHVAFDCMKNYLDALYDVTYVEGKDGGQRRSEPTMTTEFL 251
    |||||||
OY 241 CKECPKIHIIHIDRIKDDVBEDEOEHMRMWLHEREIKDKMLIEFYESPDEPKRRRPFQGS 300
    |||||||
DB 252 CKECPKIHIIHIDRIKDDVBEDEOEHMRMWLHEREIKDKMLIEFYESPDEPKRRRPFQGS 311
    |||||||
OY 301 VNSKLSIKKTLPMSLISGLTAGMLMTDAGRKLVTNTWYIGTLGCLMWTIKA 353
    |||||||
DB 312 VNSKLSIKKTLPMSLISGLTAGMLMTDAGRKLVTNTWYIGTLGCLMWTIKA 364
    |||||||
```

## RESULT 13

US-09-817-910-7

; Sequence 7, Application US/09817910

; GENERAL INFORMATION:

; APPLICANT: Meyers, Rachel A.

; APPLICANT: Macbeth, Kyle J.

; APPLICANT: Williamson, Mark

; APPLICANT: Rudolph-Owen, Laura A.

; APPLICANT: Tsai, Fong-Ying

; TITLE OF INVENTION: 46743 and 27417 NOVEL HUMAN

; FILE REFERENCE: ACYLTRANSFERASES AND USES THEREFOR

; CURRENT APPLICATION NUMBER: US/09/817,910

; PRIOR FILING DATE: 2001-03-26

; PRIOR APPLICATION NUMBER: US 60/192,092

; PRIOR FILING DATE: 2000-03-24

; NUMBER OF SEQ ID NOS: 15

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 7

; LENGTH: 364

; TYPE: PRF

; ORGANISM: Homo sapiens

; US-09-817-910-7

```
Query Match          100.0%; Score 353; DB 22; Length 364;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 353; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY 1 MRYLLPSVLLGTAPTYVLAMGWRLLSAFLPARFYQALDRLCYVQSWLFFFEYNTG 60
    |||||||
DB 12 MRYLLPSVLLGTAPTYVLAMGWRLLSAFLPARFYQALDRLCYVQSWLFFFEYNTG 71
    |||||||
OY 61 VOILLYGDLPKKNENIYYLANHOSTVDMIVADILAIQNALGHRVYLYKGLKMLPLYGC 120
    |||||||
```

```
DB 72 VOILLYGDLPKKNENIYYLANHOSTVDMIVADILAIQNALGHRVYLYKGLKMLPLYGC 131
    |||||||
OY 121 YFAOHGIIYVRSKAKFENEKEMRNKLOSVDAGTPMYLVIFPEGTRYNPEOTKVLASQAQF 180
    |||||||
DB 132 YFAOHGIIYVRSKAKFENEKEMRNKLOSVDAGTPMYLVIFPEGTRYNPEOTKVLASQAQF 191
    |||||||
OY 181 AAOGLAVLKHVLTLPRIKATHVAFDCMKNYLDALYDVTYVEGKDGGQRRSEPTMTTEFL 240
    |||||||
DB 192 AAOGLAVLKHVLTLPRIKATHVAFDCMKNYLDALYDVTYVEGKDGGQRRSEPTMTTEFL 251
    |||||||
OY 241 CKECPKIHIIHIDRIKDDVBEDEOEHMRMWLHEREIKDKMLIEFYESPDEPKRRRPFQGS 300
    |||||||
DB 252 CKECPKIHIIHIDRIKDDVBEDEOEHMRMWLHEREIKDKMLIEFYESPDEPKRRRPFQGS 311
    |||||||
OY 301 VNSKLSIKKTLPMSLISGLTAGMLMTDAGRKLVTNTWYIGTLGCLMWTIKA 353
    |||||||
DB 312 VNSKLSIKKTLPMSLISGLTAGMLMTDAGRKLVTNTWYIGTLGCLMWTIKA 364
    |||||||
```

## RESULT 14

US-09-853-526-5

; Sequence 5, Application US/09853526

; GENERAL INFORMATION:

; APPLICANT: Blumenfeld, Marta

; APPLICANT: Cohen, Daniel

; APPLICANT: Bougueleret, Lydie

; TITLE OF INVENTION: PROSTATE CANCER GENE

; FILE REFERENCE: GENSET 18CP1CP

; CURRENT APPLICATION NUMBER: US/09/853,526

; CURRENT FILING DATE: 2001-05-11

; PRIOR APPLICATION NUMBER: 09/318,907

; PRIOR FILING DATE: 1999-06-23

; PRIOR APPLICATION NUMBER: 08/996,306

; PRIOR FILING DATE: 1997-12-22

; PRIOR APPLICATION NUMBER: 60/099,658

; PRIOR FILING DATE: 1998-09-09

; PRIOR APPLICATION NUMBER: 09/218,207

; PRIOR FILING DATE: 1998-12-22

; NUMBER OF SEQ ID NOS: 578

; SOFTWARE: Patent.pm

; SEQ ID NO 5

; LENGTH: 364

; TYPE: PRF

; ORGANISM: Homo sapiens

; US-09-853-526-5

```
Query Match          100.0%; Score 353; DB 22; Length 364;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 353; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY 1 MRYLLPSVLLGTAPTYVLAMGWRLLSAFLPARFYQALDRLCYVQSWLFFFEYNTG 60
    |||||||
DB 12 MRYLLPSVLLGTAPTYVLAMGWRLLSAFLPARFYQALDRLCYVQSWLFFFEYNTG 71
    |||||||
OY 61 VOILLYGDLPKKNENIYYLANHOSTVDMIVADILAIQNALGHRVYLYKGLKMLPLYGC 120
    |||||||
DB 72 VOILLYGDLPKKNENIYYLANHOSTVDMIVADILAIQNALGHRVYLYKGLKMLPLYGC 131
    |||||||
OY 121 YFAOHGIIYVRSKAKFENEKEMRNKLOSVDAGTPMYLVIFPEGTRYNPEOTKVLASQAQF 180
    |||||||
DB 132 YFAOHGIIYVRSKAKFENEKEMRNKLOSVDAGTPMYLVIFPEGTRYNPEOTKVLASQAQF 191
    |||||||
OY 181 AAOGLAVLKHVLTLPRIKATHVAFDCMKNYLDALYDVTYVEGKDGGQRRSEPTMTTEFL 240
    |||||||
DB 192 AAOGLAVLKHVLTLPRIKATHVAFDCMKNYLDALYDVTYVEGKDGGQRRSEPTMTTEFL 251
    |||||||
OY 241 CKECPKIHIIHIDRIKDDVBEDEOEHMRMWLHEREIKDKMLIEFYESPDEPKRRRPFQGS 300
    |||||||
DB 252 CKECPKIHIIHIDRIKDDVBEDEOEHMRMWLHEREIKDKMLIEFYESPDEPKRRRPFQGS 311
    |||||||
OY 301 VNSKLSIKKTLPMSLISGLTAGMLMTDAGRKLVTNTWYIGTLGCLMWTIKA 353
    |||||||
```

Db 312 VNSKLSIKTKLPMSLLSGLTAGMLMTDAGRKLYVNTWYIGTLLGCLWYTIKA 364

## RESULT 15

US-09-901-484A-5  
; Sequence 5, Application US/09901484A  
; GENERAL INFORMATION:  
; APPLICANT: Cohen, Daniel  
; APPLICANT: Blumenfeld, Marta  
; APPLICANT: Chumakov, Ilya  
; APPLICANT: Bouguelerel, Lydie  
; TITLE OF INVENTION: Prostate Cancer Gene  
; FILE REFERENCE: GEN-T11XC3D2  
; CURRENT APPLICATION NUMBER: US/09/901,484A  
; CURRENT FILING DATE: 2001-07-09  
; PRIOR APPLICATION NUMBER: US 08/996,306  
; PRIOR FILING DATE: 1997-12-22  
; PRIOR APPLICATION NUMBER: US 60/099,658  
; PRIOR FILING DATE: 1998-09-09  
; PRIOR APPLICATION NUMBER: US 09/218,207  
; PRIOR FILING DATE: 1998-12-22  
; PRIOR APPLICATION NUMBER: US 09/338,907  
; PRIOR FILING DATE: 1999-06-23  
; PRIOR APPLICATION NUMBER: US 09/853,526  
; PRIOR FILING DATE: 2001-05-11  
; NUMBER OF SEQ ID NOS: 578  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 5  
; LENGTH: 364  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-901-484A-5

Query Match 100.0%; Score 353; DB 23; Length 364;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 353; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MRYLLPSVYLLGTAPRYVYLANGVWRLLSAFLPARFYQALDDRLXYQOSMVLFFFEENTYG 60  
DB 12 MRYLLPSVYLLGTAPRYVYLANGVWRLLSAFLPARFYQALDDRLXYQOSMVLFFFEENTYG 71  
QY 61 VOILLYGDLPRKNKENIYLANHOSYVDWIVADILAIQNALGHVRYVLEKGLKMLPLYGC 120  
DB 72 VOILLYGDLPRKNKENIYLANHOSYVDWIVADILAIQNALGHVRYVLEKGLKMLPLYGC 131  
QY 121 YFAQHGIGIYVKSANFNEKEMRNKLQSYVDAGTPLYLIFPEGTRYNPQOTKYLASQAF 180  
DB 132 YFAQHGIGIYVKSANFNEKEMRNKLQSYVDAGTPLYLIFPEGTRYNPQOTKYLASQAF 191  
QY 181 AAQRGLAVLAKHVLTPRIKATHAFCQMKNYLDIAIVYVYBEKDGQGRRESPTTEFL 240  
DB 192 AAQRGLAVLAKHVLTPRIKATHAFCQMKNYLDIAIVYVYBEKDGQGRRESPTTEFL 251  
QY 241 CKECPKIHIDRIKDVPEOEHRMRLHERFEIKDKMLIEFYESPDERKRRPGRKS 300  
DB 252 CKECPKIHIDRIKDVPEOEHRMRLHERFEIKDKMLIEFYESPDERKRRPGRKS 311  
QY 301 VNSKLSIKTKLPMSLLSGLTAGMLMTDAGRKLYVNTWYIGTLLGCLWYTIKA 353  
DB 312 VNSKLSIKTKLPMSLLSGLTAGMLMTDAGRKLYVNTWYIGTLLGCLWYTIKA 364

Search completed: August 28, 2002, 11:25:19  
Job time: 497 sec

**THIS PAGE BLANK (USPTO)**

GenCore version 4.5  
Copyright (c) 1993 - 2000 Comugen Ltd.

OM protein - protein search, using sw model

Run on: August 28, 2002, 11:17:27 ; Search time 83.88 Seconds  
(without alignments)  
1018.645 Million cell updates/sec

Title: US-09-853-526-4

Perfect score: 353  
Sequence: 1 MRYLPSVVLGTAFTYVLA.....YVNTWYTGILGIMWTIKA 353

Scoring table: OLIGO  
Gapop 60.0 , Gapext 60.0

Searched: 747981 seqs, 242050750 residues

Word size : 0

Total number of hits satisfying chosen parameters: 747981

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database :

Pending\_Patents\_AA\_New:\*  
1: /cgn2\_6/ptodata/1/paa/PCT\_NEW\_COMB.pep:\*  
2: /cgn2\_6/ptodata/1/paa/US06\_NEW\_COMB.pep:\*  
3: /cgn2\_6/ptodata/1/paa/US07\_NEW\_COMB.pep:\*  
4: /cgn2\_6/ptodata/1/paa/US08\_NEW\_COMB.pep:\*  
5: /cgn2\_6/ptodata/1/paa/US09\_NEW\_COMB.pep:\*  
6: /cgn2\_6/ptodata/1/paa/US10\_NEW\_COMB.pep:\*  
7: /cgn2\_6/ptodata/1/paa/US60\_NEW\_COMB.pep:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	353	100.0	364	US-10-184-648-63	Sequence 63, Appl
2	269	76.2	269	US-10-074-045-47	Sequence 47, Appl
3	252	71.4	353	US-09-629-469A-13028	Sequence 13028, A
4	9	2.5	135	US-10-184-648-52	Sequence 52, Appl
5	9	2.5	157	US-10-184-648-66	Sequence 66, Appl
6	9	2.5	240	US-09-895-913A-184	Sequence 184, App
7	8	2.3	236	US-09-791-537-5820	Sequence 1538, App
8	8	2.3	459	US-09-791-537-110883	Sequence 5820, Ap
9	8	2.3	483	US-09-791-537-110642	Sequence 110642, Sequence
10	8	2.3	44	US-09-826-734A-142	Sequence 142, Appl
11	7	2.0	77	US-09-548-936C-17	Sequence 17, Appl
12	7	2.0	96	US-09-791-537-150979	Sequence 150979, Sequence
13	7	2.0	117	US-09-791-537-31675	Sequence 31675, A
14	7	2.0	126	US-09-791-537-63528	Sequence 63528, A
15	7	2.0	194	US-09-791-537-53250	Sequence 53250, A
16	7	2.0	195	US-10-164-966-9	Sequence 9, Appl
17	7	2.0	195	US-10-184-648-51	Sequence 51, Appl
18	7	2.0	195	US-10-184-648-67	Sequence 67, Appl
19	7	2.0	195	US-10-184-648-86	Sequence 86, Appl
20	7	2.0	208	PCT-US02-10824-132	Sequence 132, App
21	7	2.0	208	US-09-791-537-57367	Sequence 57367, A
22	7	2.0	208	US-09-791-537-99121	Sequence 99121, A
23	7	2.0	208	US-09-791-537-128069	Sequence 128069, Sequence
24	7	2.0	208	US-09-791-537-142901	Sequence 142901, Sequence
25	7	2.0	208	US-09-791-537-143774	Sequence 143774, Sequence
26	7	2.0	208	US-09-791-537-143774	Sequence 143774, Sequence

27	7	2.0	208	US-10-096-327-2	Sequence 2, Appl
28	7	2.0	208	US-10-179-131-7124	Sequence 7124, Ap
29	7	2.0	208	US-10-189-360-5	Sequence 5, Appl
30	7	2.0	208	US-10-189-360-7	Sequence 7, Appl
31	7	2.0	208	US-10-189-360-8	Sequence 8, Appl
32	7	2.0	208	US-10-138-158-18	Sequence 18, Appl
33	7	2.0	234	US-09-791-537-148811	Sequence 148811, Sequence
34	7	2.0	246	US-09-791-537-43519	Sequence 43519, A
35	7	2.0	250	US-09-791-537-67119	Sequence 67119, A
36	7	2.0	251	US-09-791-537-29660	Sequence 29660, A
37	7	2.0	253	US-10-135-881-22428	Sequence 22428, A
38	7	2.0	262	US-09-791-537-751	Sequence 751, App
39	7	2.0	263	US-10-155-881-18654	Sequence 18654, A
40	7	2.0	264	US-09-791-537-24942	Sequence 24942, A
41	7	2.0	269	US-10-053-853A-1244	Sequence 1244, Ap
42	7	2.0	270	US-10-211-364-1026	Sequence 1026, Ap
43	7	2.0	286	US-60-360-039-1733	Sequence 1733, Ap
44	7	2.0	311	US-60-360-039-22196	Sequence 22196, A
45	7	2.0	364	US-09-791-537-42339	Sequence 42339, A

#### ALIGNMENTS

RESULT 1  
US-10-184-648-63  
Sequence 63, Application US/10184648  
GENERAL INFORMATION:  
APPLICANT: Meyers, Rachel E.  
APPLICANT: Williamson, Mark  
APPLICANT: Tsai, Feng-Ying  
APPLICANT: Hunter, John J.  
APPLICANT: Macbeth, Kyle J.  
APPLICANT: Rudolph-Owen, Laura A.  
APPLICANT: Leiby, Kevin R.  
APPLICANT: Kapeller-Libermann, Rosana  
APPLICANT: Olandt, Peter J.  
TITLE OF INVENTION: NOVEL HUMAN TRANSFERASE FAMILY MEMBERS AND USES THEREOF  
FILE REFERENCE: 10448-192001  
CURRENT APPLICATION NUMBER: US/10/184,648  
CURRENT FILING DATE: 2002-06-27  
PRIOR APPLICATION NUMBER: US 09/815,028  
PRIOR FILING DATE: 2001-03-22  
PRIOR APPLICATION NUMBER: PCT/US01/09358  
PRIOR FILING DATE: 2001-03-22  
PRIOR APPLICATION NUMBER: US 60/191,964  
PRIOR FILING DATE: 2000-03-24  
PRIOR APPLICATION NUMBER: US 09/801,220  
PRIOR FILING DATE: 2001-03-07  
PRIOR APPLICATION NUMBER: PCT/US01/07269  
PRIOR FILING DATE: 2001-03-07  
PRIOR APPLICATION NUMBER: US 60/187,456  
PRIOR FILING DATE: 2000-03-07  
PRIOR APPLICATION NUMBER: US 09/816,714  
PRIOR FILING DATE: 2001-03-23  
PRIOR APPLICATION NUMBER: PCT/US01/09468  
PRIOR FILING DATE: 2001-03-23  
PRIOR APPLICATION NUMBER: US 60/191,865  
PRIOR FILING DATE: 2000-03-24  
PRIOR APPLICATION NUMBER: US 09/844,948  
PRIOR FILING DATE: 2001-04-27  
PRIOR APPLICATION NUMBER: PCT/US01/13805  
PRIOR FILING DATE: 2001-04-27  
PRIOR APPLICATION NUMBER: US 60/200,604  
PRIOR FILING DATE: 2000-04-28  
PRIOR APPLICATION NUMBER: US 09/861,164  
PRIOR FILING DATE: 2001-05-18  
PRIOR APPLICATION NUMBER: PCT/US01/16292  
PRIOR FILING DATE: 2001-05-18  
PRIOR APPLICATION NUMBER: US 60/205,408  
PRIOR FILING DATE: 2000-05-19  
PRIOR APPLICATION NUMBER: US 09/883,060  
PRIOR FILING DATE: 2001-06-15

```

; PRIOR APPLICATION NUMBER: PCT/US01/19138
; PRIOR FILING DATE: 2001-06-15
; PRIOR APPLICATION NUMBER: US 60/212,079
; PRIOR FILING DATE: 2000-06-15
; PRIOR APPLICATION NUMBER: US 09/962,678
; PRIOR FILING DATE: 2001-09-25
; PRIOR APPLICATION NUMBER: PCT/US01/29963
; PRIOR FILING DATE: 2001-09-25
; PRIOR APPLICATION NUMBER: US 60/235,044
; PRIOR FILING DATE: 2000-09-25
; PRIOR APPLICATION NUMBER: US 09/973,457
; PRIOR FILING DATE: 2001-10-09
; PRIOR APPLICATION NUMBER: US 60/238,849
; PRIOR FILING DATE: 2000-10-06
; PRIOR APPLICATION NUMBER: US 10/072,285
; PRIOR FILING DATE: 2002-02-08
; PRIOR APPLICATION NUMBER: PCT/US02/03736
; PRIOR FILING DATE: 2002-02-08
; PRIOR APPLICATION NUMBER: US 60/267,494
; PRIOR FILING DATE: 2001-02-08
; PRIOR APPLICATION NUMBER: US 09/817,910
; PRIOR FILING DATE: 2001-03-26
; PRIOR APPLICATION NUMBER: PCT/US01/09633
; PRIOR FILING DATE: 2001-03-26
; PRIOR APPLICATION NUMBER: US 60/192,092
; PRIOR FILING DATE: 2000-03-24
; PRIOR APPLICATION NUMBER: US 09/842,528
; PRIOR FILING DATE: 2001-04-25
; PRIOR APPLICATION NUMBER: PCT/US01/40607
; PRIOR FILING DATE: 2001-04-25
; PRIOR APPLICATION NUMBER: US 60/199,500
; PRIOR FILING DATE: 2000-04-25
; PRIOR APPLICATION NUMBER: US 09/882,836
; PRIOR FILING DATE: 2001-06-15
; PRIOR APPLICATION NUMBER: PCT/US01/19543
; PRIOR FILING DATE: 2001-06-15
; PRIOR APPLICATION NUMBER: US 60/211,730
; PRIOR FILING DATE: 2000-06-15
; PRIOR APPLICATION NUMBER: US 09/882,872
; PRIOR FILING DATE: 2001-06-15
; PRIOR APPLICATION NUMBER: PCT/US01/19153
; PRIOR FILING DATE: 2001-06-15
; PRIOR APPLICATION NUMBER: US 60/212,077
; PRIOR FILING DATE: 2000-06-15
; NUMBER OF SEQ ID NOS: 90
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 63
; LENGTH: 364
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-184-648-63

```

```

Query Match          100.0%; Score 353; DB 6; Length 364;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 353; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1 MRYLLPSVLLGTAPTYVLAAGVWRLSAPLPARYQALDRLCYVQSMVLFEEENTG 60
    |||||
DB 12 MRYLLPSVLLGTAPTYVLAAGVWRLSAPLPARYQALDRLCYVQSMVLFEEENTG 71
    |||||
QY 61 VOILLGDLPRKNENIITLANHSTVDWIVADILAIQNALGHRVYVLEGLKMLPLYGC 120
    |||||
DB 72 VOILLGDLPRKNENIITLANHSTVDWIVADILAIQNALGHRVYVLEGLKMLPLYGC 131
    |||||
QY 121 YFAHQGIIYKRSKAFNEKEKRNKLSQSVDAQTPMYLVIPEEGTRYNPEQTKVLSASQAF 180
    |||||
DB 132 YFAHQGIIYKRSKAFNEKEKRNKLSQSVDAQTPMYLVIPEEGTRYNPEQTKVLSASQAF 191
    |||||
QY 181 AAQGLAVLKHLVLPRIKATHVAFDCKMKNYDAIYDTVYVYEGKDDGGQRESPTMTEFL 240
    |||||
DB 192 AAQGLAVLKHLVLPRIKATHVAFDCKMKNYDAIYDTVYVYEGKDDGGQRESPTMTEFL 251
    |||||

```

```

QY 241 CKECPKIHIDRIDDKDVPEOEHRMRWLHERREIKDKMLIEFYESPDPERRRRFPCKS 300
    |||||
DB 252 CKECPKIHIDRIDDKDVPEOEHRMRWLHERREIKDKMLIEFYESPDPERRRRFPCKS 311
    |||||
QY 301 VNSKSIKTKTLPMSILISGLTAGMLMTDAGRKLYVNTIYGTLLGCLWVTIKA 353
    |||||
DB 312 VNSKSIKTKTLPMSILISGLTAGMLMTDAGRKLYVNTIYGTLLGCLWVTIKA 364
    |||||

```

```

RESULT 2
US-10-074-045-47
; Sequence 47, Application US/10074045
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PT221C1
; CURRENT APPLICATION NUMBER: US/10/074,045
; CURRENT FILING DATE: 2002-02-14
; Prior Application removed - See file wrapper or Palm
; NUMBER OF SEQ ID NOS: 74
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 47
; LENGTH: 269
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-074-045-47

```

```

Query Match          76.2%; Score 269; DB 6; Length 269;
Best Local Similarity 100.0%; Pred. No. 8,7e+268;
Matches 269; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 85 TVDWIVADILAIQNALGHRVYVLEGLKMLPLGCFYAGHGIIYKRSKAFNEKRNK 144
    |||||
DB 1 TVDWIVADILAIQNALGHRVYVLEGLKMLPLGCFYAGHGIIYKRSKAFNEKRNK 60
    |||||
QY 145 LQSYVDAGTPMYLVIPEEGTRYNPEQTKVLSASQAFPAQGLAVLKHLVLPRIKATHVAF 204
    |||||
DB 61 LQSYVDAGTPMYLVIPEEGTRYNPEQTKVLSASQAFPAQGLAVLKHLVLPRIKATHVAF 120
    |||||
QY 205 DCKMKNYDAIYDTVYVYEGKDDGGQRESPTMTEFLCKECPKIHIDRIDDKVPEOE 264
    |||||
DB 121 DCKMKNYDAIYDTVYVYEGKDDGGQRESPTMTEFLCKECPKIHIDRIDDKVPEOE 180
    |||||
QY 265 HMRWLHERREIKDKMLIEFYESPDPERRRRFPCKSVNSKSIKTKTLPMSILISGLTAGM 324
    |||||
DB 181 HMRWLHERREIKDKMLIEFYESPDPERRRRFPCKSVNSKSIKTKTLPMSILISGLTAGM 240
    |||||
QY 325 LMTDAGRKLYVNTIYGTLLGCLWVTIKA 353
    |||||
DB 241 LMTDAGRKLYVNTIYGTLLGCLWVTIKA 269
    |||||

```

```

RESULT 3
US-09-629-469A-13028
; Sequence 13028, Application US/09629469A
; GENERAL INFORMATION:
; APPLICANT: OTA, TOSHIO
; APPLICANT: ISOGAI, TAKAO
; APPLICANT: NISHIKAWA, TETSUO
; APPLICANT: HAYASHI, KOJI
; APPLICANT: SAITO, KAORU
; APPLICANT: YAMAMOTO, JUNICHI
; APPLICANT: ISHII, SHIZUKO
; APPLICANT: SUGIYAMA, TOMOYASU
; APPLICANT: WAKAMATSU, AI
; APPLICANT: NAGAI, KEIICHI
; APPLICANT: OTSUKI, TETSUJI
; TITLE OF INVENTION: PRIMERS FOR SYNTHESIZING FULL-LENGTH CDNA AND THEIR USE
; FILE REFERENCE: 084335/0123
; CURRENT APPLICATION NUMBER: US/09/629,469A
; CURRENT FILING DATE: 2000-07-28
; PRIOR APPLICATION NUMBER: JP 1999-248036

```



;; PRIOR FILING DATE: 1999-07-29  
;; PRIOR APPLICATION NUMBER: JP 1999-300253  
;; PRIOR FILING DATE: 1999-08-27  
;; PRIOR APPLICATION NUMBER: JP 2000-118776  
;; PRIOR FILING DATE: 2000-01-11  
;; PRIOR APPLICATION NUMBER: JP 2000-183767  
;; PRIOR FILING DATE: 2000-05-02  
;; PRIOR APPLICATION NUMBER: JP 2000-241899  
;; PRIOR FILING DATE: 2000-06-09  
;; PRIOR APPLICATION NUMBER: 60/159,590  
;; PRIOR FILING DATE: 1999-10-18  
;; PRIOR APPLICATION NUMBER: 60/183,322  
;; PRIOR FILING DATE: 2000-02-17  
;; NUMBER OF SEQ ID NOS: 19025  
;; SOFTWARE: PatentIn Ver. 2.1  
;; SEQ ID NO 13028  
;; LENGTH: 353  
;; TYPE: PRT  
;; ORGANISM: Homo sapiens  
US-09-629-469A-13028

Query Match 71.4%; Score 252; DB 5; Length 353;  
Best Local Similarity 99.7%; Pred. No. 3.3e-250;  
Matches 352; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MRLLSPSVLLGTAFTYVLAAGWRLLSAFLPARFYQALDDRLCYQSVLFEFFENYTG 60  
DB 1 MRLLSPSVLLGTAFTYVLAAGWRLLSAFLPARFYQALDDRLCYQSVLFEFFENYTG 60  
QY 61 VQILLXGDLPRKNENIYVLANHSTVDWIVADILAIKRONALGHVRYLKGKMLPLYGC 120  
DB 61 VQILLXGDLPRKNENIYVLANHSTVDWIVADILAIKRONALGHVRYLKGKMLPLYGC 120  
QY 121 YFAOHGIIYKRSAKFNEKEMRNKLSQSYVDAGTPMYLVIFPEGTRYNPEQTKVLSAQAF 180  
DB 121 YFAOHGIIYKRSAKFNEKEMRNKLSQSYVDAGTPMYLVIFPEGTRYNPEQTKVLSAQAF 180  
QY 181 AARGLAIVLKHVLTPIKATHVAFDCMKNYLDIYDVTYVEKDKDGGQRRSEPTMEFL 240  
DB 181 AARGLAIVLKHVLTPIKATHVAFDCMKNYLDIYDVTYVEKDKDGGQRRSEPTMEFL 240  
QY 241 CKCEPKIHIDRIDKKDVEEEOEHMRWMLHERFEIKDKMLIEFYESPDERKKRFPFGKS 300  
DB 241 CKCEPKIHIDRIDKKDVEEEOEHMRWMLHERFEIKDKMLIEFYESPDERKKRFPFGKS 300  
QY 301 VNSKLSIKTLPMSLLISGLTAGMLMTDAGRKLIVNTWIIYGTLLGLMTWITKA 353  
DB 301 VNSKLSIKTLPMSLLISGLTAGMLMTDAGRKLIVNTWIIYGTLLGLMTWITKA 353

RESULT 4  
US-10-184-648-52  
;; Sequence 52, Application US/10184648  
;; GENERAL INFORMATION:  
;; APPLICANT: Meyers, Rachel E.  
;; APPLICANT: Williamson, Mark  
;; APPLICANT: Tsai, Fong-Ying  
;; APPLICANT: Hunter, John J.  
;; APPLICANT: Macbeth, Kyle J.  
;; APPLICANT: Rudolph-Owen, Laura A.  
;; APPLICANT: Leiby, Kevin R.  
;; APPLICANT: Kapeller-libermann, Rosana  
;; APPLICANT: Olandt, Peter J.  
;; TITLE OF INVENTION: NOVEL HUMAN TRANSFERASE FAMILY MEMBERS AND USES THEREOF  
;; FILE REFERENCE: 10448-192001  
;; CURRENT APPLICATION NUMBER: US/10/184,648  
;; CURRENT FILING DATE: 2002-06-27  
;; PRIOR APPLICATION NUMBER: US 09/815,028  
;; PRIOR FILING DATE: 2001-03-22  
;; PRIOR APPLICATION NUMBER: PCT/US01/09358  
;; PRIOR FILING DATE: 2001-03-22  
;; PRIOR APPLICATION NUMBER: US 60/191,964

;; PRIOR FILING DATE: 2000-03-24  
;; PRIOR APPLICATION NUMBER: US 09/801,220  
;; PRIOR FILING DATE: 2001-03-07  
;; PRIOR APPLICATION NUMBER: PCT/US01/07269  
;; PRIOR FILING DATE: 2001-03-07  
;; PRIOR APPLICATION NUMBER: US 60/187,456  
;; PRIOR FILING DATE: 2000-03-07  
;; PRIOR APPLICATION NUMBER: US 09/816,714  
;; PRIOR FILING DATE: 2001-03-23  
;; PRIOR APPLICATION NUMBER: PCT/US01/09468  
;; PRIOR FILING DATE: 2001-03-23  
;; PRIOR APPLICATION NUMBER: US 60/191,865  
;; PRIOR FILING DATE: 2000-03-24  
;; PRIOR APPLICATION NUMBER: US 09/844,948  
;; PRIOR FILING DATE: 2001-04-27  
;; PRIOR APPLICATION NUMBER: PCT/US01/13805  
;; PRIOR FILING DATE: 2001-04-27  
;; PRIOR APPLICATION NUMBER: US 60/200,604  
;; PRIOR FILING DATE: 2000-04-28  
;; PRIOR APPLICATION NUMBER: US 09/861,164  
;; PRIOR FILING DATE: 2001-05-18  
;; PRIOR APPLICATION NUMBER: PCT/US01/16292  
;; PRIOR FILING DATE: 2001-05-18  
;; PRIOR APPLICATION NUMBER: US 60/205,408  
;; PRIOR FILING DATE: 2000-05-19  
;; PRIOR APPLICATION NUMBER: US 09/883,060  
;; PRIOR FILING DATE: 2001-06-15  
;; PRIOR APPLICATION NUMBER: PCT/US01/19138  
;; PRIOR FILING DATE: 2001-06-15  
;; PRIOR APPLICATION NUMBER: US 60/212,079  
;; PRIOR FILING DATE: 2000-06-15  
;; PRIOR APPLICATION NUMBER: US 09/962,678  
;; PRIOR FILING DATE: 2001-09-25  
;; PRIOR APPLICATION NUMBER: PCT/US01/29963  
;; PRIOR FILING DATE: 2001-09-25  
;; PRIOR APPLICATION NUMBER: US 60/235,044  
;; PRIOR FILING DATE: 2000-09-25  
;; PRIOR APPLICATION NUMBER: US 09/973,457  
;; PRIOR FILING DATE: 2001-10-09  
;; PRIOR APPLICATION NUMBER: US 60/238,849  
;; PRIOR FILING DATE: 2000-10-06  
;; PRIOR APPLICATION NUMBER: US 10/072,285  
;; PRIOR FILING DATE: 2002-02-08  
;; PRIOR APPLICATION NUMBER: PCT/US02/03736  
;; PRIOR FILING DATE: 2002-02-08  
;; PRIOR APPLICATION NUMBER: US 60/267,494  
;; PRIOR FILING DATE: 2001-02-08  
;; PRIOR APPLICATION NUMBER: US 09/817,910  
;; PRIOR FILING DATE: 2001-03-26  
;; PRIOR APPLICATION NUMBER: PCT/US01/09633  
;; PRIOR FILING DATE: 2001-03-26  
;; PRIOR APPLICATION NUMBER: US 60/192,092  
;; PRIOR FILING DATE: 2000-03-24  
;; PRIOR APPLICATION NUMBER: US 09/842,528  
;; PRIOR FILING DATE: 2001-04-25  
;; PRIOR APPLICATION NUMBER: PCT/US01/40607  
;; PRIOR FILING DATE: 2001-04-25  
;; PRIOR APPLICATION NUMBER: US 60/199,500  
;; PRIOR FILING DATE: 2000-04-25  
;; PRIOR APPLICATION NUMBER: US 09/882,836  
;; PRIOR FILING DATE: 2001-06-15  
;; PRIOR APPLICATION NUMBER: PCT/US01/19543  
;; PRIOR FILING DATE: 2001-06-15  
;; PRIOR APPLICATION NUMBER: US 60/211,730  
;; PRIOR FILING DATE: 2000-06-15  
;; PRIOR APPLICATION NUMBER: US 09/882,872  
;; PRIOR FILING DATE: 2001-06-15  
;; PRIOR APPLICATION NUMBER: PCT/US01/19153  
;; PRIOR FILING DATE: 2001-06-15  
;; PRIOR APPLICATION NUMBER: US 60/212,077  
;; PRIOR FILING DATE: 2000-06-15  
;; NUMBER OF SEQ ID NOS: 90  
;; SOFTWARE: FastSeq for Windows Version 4.0

```
SEQ ID NO 52
LENGTH: 135
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: consensus sequence
US-10-184-648-52
```

```
Query Match          2.5%; Score 9; DB 6; Length 135;
Best Local Similarity 100.0%; Pred. No. 0 67;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY 157 LVIRPEGTR 165
      |||||
Db 85 LVIRPEGTR 93
```

```
RESULT 5
US-10-184-648-66
Sequence 66, Application US/10184648
GENERAL INFORMATION:
APPLICANT: Meyers, Rachel E.
APPLICANT: Williamson, Mark
APPLICANT: Tsai, Fong-Ying
APPLICANT: Hunter, John J.
APPLICANT: Macbeth, Kyle J.
APPLICANT: Rudolph-Owen, Laura A.
APPLICANT: Leiby, Kevin R.
APPLICANT: Kapeller-Libermann, Rosana
APPLICANT: Olandt, Peter J.
TITLE OF INVENTION: NOVEL HUMAN TRANSERASE FAMILY MEMBERS AND USES THEREOF
FILE REFERENCE: 10448-192001
CURRENT APPLICATION NUMBER: US/10/184,648
PRIOR FILING DATE: 2002-06-27
PRIOR APPLICATION NUMBER: US 09/815,028
PRIOR FILING DATE: 2001-03-22
PRIOR APPLICATION NUMBER: PCT/US01/09358
PRIOR FILING DATE: 2001-03-22
PRIOR APPLICATION NUMBER: US 60/191,964
PRIOR FILING DATE: 2000-03-24
PRIOR APPLICATION NUMBER: US 09/801,220
PRIOR FILING DATE: 2001-03-07
PRIOR APPLICATION NUMBER: PCT/US01/07269
PRIOR FILING DATE: 2001-03-07
PRIOR APPLICATION NUMBER: US 60/187,456
PRIOR FILING DATE: 2000-03-07
PRIOR APPLICATION NUMBER: US 09/816,714
PRIOR FILING DATE: 2001-03-23
PRIOR APPLICATION NUMBER: PCT/US01/09468
PRIOR FILING DATE: 2001-03-23
PRIOR APPLICATION NUMBER: US 60/191,865
PRIOR FILING DATE: 2000-03-24
PRIOR APPLICATION NUMBER: US 09/844,948
PRIOR FILING DATE: 2001-04-27
PRIOR APPLICATION NUMBER: PCT/US01/13805
PRIOR FILING DATE: 2001-04-27
PRIOR APPLICATION NUMBER: US 60/200,604
PRIOR FILING DATE: 2000-04-28
PRIOR APPLICATION NUMBER: US 09/861,164
PRIOR FILING DATE: 2001-05-18
PRIOR APPLICATION NUMBER: PCT/US01/16292
PRIOR FILING DATE: 2001-05-18
PRIOR APPLICATION NUMBER: US 60/205,408
PRIOR FILING DATE: 2000-05-19
PRIOR APPLICATION NUMBER: US 09/883,060
PRIOR FILING DATE: 2001-06-15
PRIOR APPLICATION NUMBER: PCT/US01/19138
PRIOR FILING DATE: 2001-06-15
PRIOR APPLICATION NUMBER: US 60/212,079
PRIOR FILING DATE: 2000-06-15
PRIOR APPLICATION NUMBER: US 09/962,678
PRIOR FILING DATE: 2001-09-25
```

```
PRIOR APPLICATION NUMBER: PCT/US01/29963
PRIOR FILING DATE: 2001-09-25
PRIOR APPLICATION NUMBER: US 60/235,044
PRIOR FILING DATE: 2000-09-25
PRIOR APPLICATION NUMBER: US 09/973,457
PRIOR FILING DATE: 2001-10-09
PRIOR APPLICATION NUMBER: US 60/238,849
PRIOR FILING DATE: 2000-10-06
PRIOR APPLICATION NUMBER: US 10/072,285
PRIOR FILING DATE: 2002-02-08
PRIOR APPLICATION NUMBER: PCT/US02/03736
PRIOR FILING DATE: 2002-02-08
PRIOR APPLICATION NUMBER: US 60/267,494
PRIOR FILING DATE: 2001-02-08
PRIOR APPLICATION NUMBER: US 09/817,910
PRIOR FILING DATE: 2001-03-26
PRIOR APPLICATION NUMBER: PCT/US01/09633
PRIOR FILING DATE: 2001-03-26
PRIOR APPLICATION NUMBER: US 60/192,092
PRIOR FILING DATE: 2000-03-24
PRIOR APPLICATION NUMBER: US 09/842,528
PRIOR FILING DATE: 2001-04-25
PRIOR APPLICATION NUMBER: PCT/US01/40607
PRIOR FILING DATE: 2001-04-25
PRIOR APPLICATION NUMBER: US 60/199,500
PRIOR FILING DATE: 2000-04-25
PRIOR APPLICATION NUMBER: US 09/882,836
PRIOR FILING DATE: 2001-06-15
PRIOR APPLICATION NUMBER: PCT/US01/19543
PRIOR FILING DATE: 2001-06-15
PRIOR APPLICATION NUMBER: US 60/211,730
PRIOR FILING DATE: 2000-06-15
PRIOR APPLICATION NUMBER: US 09/882,872
PRIOR FILING DATE: 2001-06-15
PRIOR APPLICATION NUMBER: PCT/US01/19153
PRIOR FILING DATE: 2001-06-15
PRIOR APPLICATION NUMBER: US 60/212,077
PRIOR FILING DATE: 2000-06-15
NUMBER OF SEQ ID NOS: 90
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 66
LENGTH: 157
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: consensus sequence
US-10-184-648-66
```

```
Query Match          2.5%; Score 9; DB 6; Length 157;
Best Local Similarity 100.0%; Pred. No. 0.77;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY 211 LDATYDVTY 219
      |||||
Db 4 LDATYDVTY 12
```

```
RESULT 6
US-09-895-913A-184
Sequence 184, Application US/09895913A
GENERAL INFORMATION:
APPLICANT: Kleantous, Harold
APPLICANT: Al-Garawi, Amal
APPLICANT: Miller, Charles
APPLICANT: Tomb, Jean Francois
APPLICANT: Oomen, Raymond P.
TITLE OF INVENTION: Identification of Polynucleotides
TITLE OF INVENTION: Encoding Novel Helicobacter Polypeptides in the Helicobacter
FILE REFERENCE: 06133/043002
CURRENT APPLICATION NUMBER: US/09/895,913A
PRIOR FILING DATE: 2001-06-29
```

; PRIOR APPLICATION NUMBER: US 08/881,227  
; PRIOR FILING DATE: 1997-06-24  
; NUMBER OF SEQ ID NOS: 368  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 184  
; LENGTH: 240  
; TYPE: PRT  
; ORGANISM: Helicobacter pylori  
US-09-895-913A-184

Query Match 2.5%; Score 9; DB 5; Length 240;  
Best Local Similarity 100.0%; Pred. No. 1.1;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 157 LVIFPEGTR 165  
Db 142 LVIFPEGTR 150

RESULT 7  
US-10-053-853A-1538  
; Sequence 1538, Application US/10053853A  
; GENERAL INFORMATION:  
; APPLICANT: HAYASHI, Hideo  
; APPLICANT: SHINAGAWA, Hideo  
; APPLICANT: MAKINO, Kozo  
; APPLICANT: HAYASHI, Tetsuya  
; APPLICANT: OHNISHI, Makoto  
; APPLICANT: HATTORI, Masahira  
; APPLICANT: KUROKAWA, Ken  
; TITLE OF INVENTION: Polynucleotide molecules and polypeptides specific to Enterohemor  
; FILE REFERENCE: 2002-00604/MMC/01704  
; CURRENT APPLICATION NUMBER: US/10/053,853A  
; CURRENT FILING DATE: 2002-05-28  
; PRIOR APPLICATION NUMBER: JP2001-112010  
; PRIOR FILING DATE: 2001-01-24  
; NUMBER OF SEQ ID NOS: 1866  
; SEQ ID NO 1538  
; LENGTH: 236  
; TYPE: PRT  
; ORGANISM: Escherichia coli O157:H7  
US-10-053-853A-1538

Query Match 2.3%; Score 8; DB 6; Length 236;  
Best Local Similarity 100.0%; Pred. No. 1.2;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 193 LTPRIKAT 200  
Db 227 LTPRIKAT 234

RESULT 8  
US-09-791-537-5820  
; Sequence 5820, Application US/09791537  
; GENERAL INFORMATION:  
; APPLICANT: Bionomix, Inc.  
; APPLICANT: Debe, Derek  
; APPLICANT: Danzer, Joseph  
; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMB  
; FILE REFERENCE: 261/210  
; CURRENT APPLICATION NUMBER: US/09/791,537  
; CURRENT FILING DATE: 2001-02-22  
; NUMBER OF SEQ ID NOS: 153055  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 5820  
; LENGTH: 262  
; TYPE: PRT  
; ORGANISM: Chlamydia trachomatis

US-09-791-537-5820

Query Match 2.3%; Score 8; DB 5; Length 262;  
Best Local Similarity 100.0%; Pred. No. 1.3;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 191 HVLPRIK 198  
Db 202 HVLPRIK 209

RESULT 9  
US-09-791-537-110883  
; Sequence 110883, Application US/09791537  
; GENERAL INFORMATION:  
; APPLICANT: Bionomix, Inc.  
; APPLICANT: Debe, Derek  
; APPLICANT: Danzer, Joseph  
; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY ME  
; FILE REFERENCE: 261/210  
; CURRENT APPLICATION NUMBER: US/09/791,537  
; CURRENT FILING DATE: 2001-02-22  
; NUMBER OF SEQ ID NOS: 153055  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 110883  
; LENGTH: 459  
; TYPE: PRT  
; ORGANISM: Bacillus subtilis  
US-09-791-537-110883

Query Match 2.3%; Score 8; DB 5; Length 459;  
Best Local Similarity 100.0%; Pred. No. 2.2;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 167 NPEOTKVL 174  
Db 292 NPEOTKVL 299

RESULT 10  
US-09-791-537-110642  
; Sequence 110642, Application US/09791537  
; GENERAL INFORMATION:  
; APPLICANT: Bionomix, Inc.  
; APPLICANT: Debe, Derek  
; APPLICANT: Danzer, Joseph  
; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY ME  
; FILE REFERENCE: 261/210  
; CURRENT APPLICATION NUMBER: US/09/791,537  
; CURRENT FILING DATE: 2001-02-22  
; NUMBER OF SEQ ID NOS: 153055  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 110642  
; LENGTH: 483  
; TYPE: PRT  
; ORGANISM: Pisum sativum  
US-09-791-537-110642

Query Match 2.3%; Score 8; DB 5; Length 483;  
Best Local Similarity 100.0%; Pred. No. 2.4;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 151 AGTPMTLV 158  
Db 184 AGTPMTLV 191

RESULT 11

```
US-09-826-734A-142
; Sequence 142, Application US/09826734A
; GENERAL INFORMATION:
; APPLICANT: Fernandes, Elma R
; APPLICANT: Mishra, Vishnu S
; APPLICANT: Leach, Martin D
; APPLICANT: Shinkets, Richard A
; APPLICANT: Zerhusen, Bryan D
; TITLE OF INVENTION: Novel Polynucleotides And Polypeptides Encoded Thereby
; FILE REFERENCE: 15966-754
; CURRENT APPLICATION NUMBER: US/09/826,734A
; CURRENT FILING DATE: 2001-04-05
; PRIOR APPLICATION NUMBER: 60/195,576
; PRIOR FILING DATE: 2000-04-26
; NUMBER OF SEQ ID NOS: 264
; SOFTWARE: Patentln Ver. 2.1
; SEQ ID NO 142
; LENGTH: 44
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-826-734A-142
```

```
Query Match          2.0%; Score 7; DB 5; Length 44;
Best Local Similarity 100.0%; Pred. No. 26;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY 159 IFEPECTR 165
      |||||||
      3 IFEPECTR 9
```

```
RESULT 12
; Sequence 17, Application US/09548936C
; GENERAL INFORMATION:
; APPLICANT: Gopez, Leonel Jorge
; APPLICANT: Saraz, Jan
; APPLICANT: Claesson-Welsh, Lena
; APPLICANT: Heldin, Carl-Henrik
; TITLE OF INVENTION: PTPL1 BINDING AGENTS
; FILE REFERENCE: L0461/7084
; CURRENT APPLICATION NUMBER: US/09/548,936C
; CURRENT FILING DATE: 2000-04-13
; PRIOR APPLICATION NUMBER: US 09/100,804
; PRIOR FILING DATE: 1998-06-19
; PRIOR APPLICATION NUMBER: US 08/596,291
; PRIOR FILING DATE: 1994-09-01
; PRIOR APPLICATION NUMBER: US 08/115,573
; NUMBER OF SEQ ID NOS: 35
; SOFTWARE: Patentln Version 3.0
; SEQ ID NO 17
; LENGTH: 77
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-548-936C-17
```

```
Query Match          2.0%; Score 7; DB 5; Length 77;
Best Local Similarity 100.0%; Pred. No. 45;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY 125 HGGIYVK 131
      |||||||
      20 HGGIYVK 26
```

```
RESULT 13
US-09-791-537-150979
; Sequence 150979, Application US/09791537
; GENERAL INFORMATION:
```

```
; APPLICANT: Bionomix, Inc.
; APPLICANT: Debe, Derek
; APPLICANT: Danzer, Joseph
; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY ME
; TITLE OF INVENTION: METHODS OF USE THEREOF
; FILE REFERENCE: 261/210
; CURRENT APPLICATION NUMBER: US/09/791,537
; CURRENT FILING DATE: 2001-02-22
; NUMBER OF SEQ ID NOS: 153055
; SOFTWARE: Patentln version 3.0
; SEQ ID NO 150979
; LENGTH: 96
; TYPE: PRT
; ORGANISM: pdb 3PDZA
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(1)
; OTHER INFORMATION: X is an unknown amino acid
US-09-791-537-150979
```

```
Query Match          2.0%; Score 7; DB 5; Length 96;
Best Local Similarity 100.0%; Pred. No. 55;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY 125 HGGIYVK 131
      |||||||
      32 HGGIYVK 38
```

```
RESULT 14
US-09-791-537-31675
; Sequence 31675, Application US/09791537
; GENERAL INFORMATION:
; APPLICANT: Bionomix, Inc.
; APPLICANT: Danzer, Joseph
; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY ME
; TITLE OF INVENTION: METHODS OF USE THEREOF
; FILE REFERENCE: 261/210
; CURRENT APPLICATION NUMBER: US/09/791,537
; CURRENT FILING DATE: 2001-02-22
; NUMBER OF SEQ ID NOS: 153055
; SOFTWARE: Patentln version 3.0
; SEQ ID NO 31675
; LENGTH: 117
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-791-537-31675
```

```
Query Match          2.0%; Score 7; DB 5; Length 117;
Best Local Similarity 100.0%; Pred. No. 66;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY 125 HGGIYVK 131
      |||||||
      43 HGGIYVK 49
```

```
RESULT 15
US-09-791-537-63528
; Sequence 63528, Application US/09791537
; GENERAL INFORMATION:
; APPLICANT: Bionomix, Inc.
; APPLICANT: Debe, Derek
; APPLICANT: Danzer, Joseph
; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY ME
; TITLE OF INVENTION: METHODS OF USE THEREOF
; FILE REFERENCE: 261/210
; CURRENT APPLICATION NUMBER: US/09/791,537
; CURRENT FILING DATE: 2001-02-22
; NUMBER OF SEQ ID NOS: 153055
```

; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 63528  
; LENGTH: 126  
; TYPE: PRT  
; ORGANISM: Mus musculus  
US-09-791-537-63528

Query Match 2.0%; Score 7; DB 5; Length 126;  
Best Local Similarity 100.0%; Pred. No. 71;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 125 HGGIYVK 131  
| | | | |  
Db 52 HGGIYVK 58

Search completed: August 28, 2002, 11:26:50  
Job time: 563 sec

**THIS PAGE BLANK (USPTO)**

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 28, 2002, 11:15:47 ; Search time 39.16 Seconds

(without alignments)  
866.178 Million cell updates/sec

Title: US-09-853-526-4

Perfect score: 353

Sequence: 1 MRYLPSVVLGTAPRYVLA.....YVMTWIXFTLLGLMTVTKA 353

Scoring table: OLIGO

Searched: 283138 seqs, 96089334 residues

Word size: 0

Total number of hits satisfying chosen parameters: 283138

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database: PIR\_71.\*

1: pir1.\*  
2: pir2.\*  
3: pir3.\*  
4: pir4.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	9	2.5	237	2	B71827
2	9	2.5	240	2	D64688
3	9	2.5	264	2	E82903
4	9	2.5	918	2	T34057
5	8	2.3	83	1	W8BPG7
6	8	2.3	83	1	W8BPG7
7	8	2.3	236	2	H90892
8	8	2.3	236	2	B87524
9	8	2.3	239	2	A85725
10	8	2.3	262	2	A71478
11	8	2.3	282	2	G71334
12	8	2.3	396	2	SS4641
13	8	2.3	411	2	C87586
14	8	2.3	424	2	T49607
15	8	2.3	458	2	G83735
16	8	2.3	459	2	E63793
17	8	2.3	483	2	T06459
18	8	2.3	611	2	A85000
19	7	2.0	54	2	SS8122
20	7	2.0	57	2	D86108
21	7	2.0	117	2	I81209
22	7	2.0	126	2	I81210
23	7	2.0	141	2	SS8117
24	7	2.0	143	2	B90523
25	7	2.0	148	2	C86732
26	7	2.0	194	2	PC1136
27	7	2.0	195	2	C83854
28	7	2.0	208	1	A38432
29	7	2.0	208	1	A41914

30	7	2.0	208	1	JC1409	heparin-binding Eg
31	7	2.0	209	2	B83329	probable acyltrans
32	7	2.0	210	2	A10342	probable acyltrans
33	7	2.0	211	2	E70476	2-acylglycerophos
34	7	2.0	225	2	H95244	conserved hypothet
35	7	2.0	226	2	E98109	conserved hypothet
36	7	2.0	230	2	E81397	probable 1-acylgly
37	7	2.0	240	2	S75162	hypothetical prote
38	7	2.0	241	2	B97019	1-acyl-sn-glycerol
39	7	2.0	246	2	T46446	hypothetical prote
40	7	2.0	247	2	A81957	1-acylglycerol-3-p
41	7	2.0	247	2	G81013	1-acyl-sn-glycerol
42	7	2.0	250	1	A31757	homeotic protein H
43	7	2.0	250	2	T27772	26S proteasome reg
44	7	2.0	250	2	E70104	1-acylglycerol-3-p
45	7	2.0	250	2	AF3384	outer membrane pro

#### ALIGNMENTS

RESULT 1  
B71827  
probable 1-acylglycerol-3-phosphate O-acyltransferase (EC 2.3.1.51) - Helicobacter pylori  
C:Species: Helicobacter pylori  
A:Variate: strain J99  
C>Date: 12-Feb-1999 #sequence\_revision 12-Feb-1999 #text\_change 06-Oct-2000  
C:Accession: B71827  
R:Alm, R.A.; Ling, L.S.L.; Molr, D.T.; King, B.L.; Brown, E.D.; Doly, P.C.; Smith, D.; Ives, C.; Gibson, R.; Merberg, D.; Mills, S.D.; Jiang, O.; Taylor, D.E.; Vovis, G.F. Nature 397, 176-180, 1999  
A>Title: Genomic sequence comparison of two unrelated isolates of the human gastric p  
A:Reference number: A71800; MID:99120557  
A:Accession: B71827  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-237 <ARN>  
A:Cross-references: GB:AE001550; GB:AE001439; NID:g4155872; PIDN:AAD06852.1; PID:g415  
A:Experimental source: strain J99  
C:Genetics:  
A:Gene: plsc  
C:Superfamily: mouse 1-acylglycerol-3-phosphate O-acyltransferase  
C:Keywords: acyltransferase; coenzyme A

Query Match 2.5%; Score 9; DB 2; Length 237;  
Best Local Similarity 100.0%; Pred. No. 0.39;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 157 LVTFPEGTR 165  
Db 142 LVTFPEGTR 150

RESULT 2  
D64688  
probable 1-acylglycerol-3-phosphate O-acyltransferase (EC 2.3.1.51) - Helicobacter pylori  
C:Species: Helicobacter pylori  
C>Date: 09-Aug-1997 #sequence\_revision 09-Aug-1997 #text\_change 06-Oct-2000  
C:Accession: D64688  
R:Tomb, J.F.; White, O.; Kariavase, A.R.; Clayton, R.A.; Sutton, G.G.; Fleischmann, R. Peterson, S.; Loftus, B.; Richardson, D.; Dodson, R.; Khalak, H.G.; Glodek, A.; McKee son, J.D.; Kelley, J.M.; Cotton, M.D.; Weidman, J.M.; Fujii, C.; Bowman, C.; Watthey, Nature 388, 539-547, 1997  
A:Authors: Wallin, E.; Hayes, W.S.; Borodovsky, M.; Karp, P.D.; Smith, H.O.; Fraser, A>Title: The complete genome sequence of the gastric pathogen Helicobacter pylori.  
A:Reference number: A64520; MID:97394467  
A:Accession: D64688  
A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-240 <TOM>  
A:Cross-references: GB:AE000636; GB:AE000511; NID:g2314517; PIDN:AAD08393.1; PID:g231  
C:Superfamily: mouse 1-acylglycerol-3-phosphate O-acyltransferase

C:Keywords: acyltransferase; coenzyme A

Query Match 2.5%; Score 9; DB 2; Length 240;  
Best Local Similarity 100.0%; Pred. No. 0.39;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 157 LVIFPECTR 165  
DB 142 LVIFPECTR 150

RESULT 3  
E82903  
1-acyl-sn-glycerol-3-phosphate acyltransferase U0344 [Imported] - Ureaplasma urealyticum  
C:Species: Ureaplasma urealyticum  
C>Date: 18-Aug-2000 #sequence\_revision 20-Aug-2000 #text\_change 20-Aug-2000  
C:Accession: E82903  
R:Glas, J.I.; Lefkowitz, E.J.; Glass, J.S.; Helner, C.R.; Chen, E.Y.; Cassell, G.H.  
submitted to GenBank, February 2000  
A:Description: The complete sequence of Ureaplasma urealyticum: Alternate views of a min  
A:Reference number: A82870  
A:Accession: E82903  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-264 <GLA>  
A:Cross-references: GB:AE002131; GB:AF222894; NID:g6899316; PIDN:AAF30753.1; GSPDB:GN001  
A:Experimental source: serovar 3; biovar 1  
C:Genetics:  
A:Gene: pISC: U0344  
A:Genetic code: SGC3

Query Match 2.5%; Score 9; DB 2; Length 264;  
Best Local Similarity 100.0%; Pred. No. 0.43;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 157 LVIFPECTR 165  
DB 156 LVIFPECTR 164

RESULT 4  
T34057  
hypothetical protein F28B3.5 - Caenorhabditis elegans  
C:Species: Caenorhabditis elegans  
C>Date: 29-Oct-1999 #sequence\_revision 29-Oct-1999 #text\_change 29-Oct-1999  
C:Accession: T34057  
R:Geisel, C.; Kramer, J.; Smith, A.  
submitted to the EMBL Data Library, May 1997  
A:Description: The sequence of C. elegans cosmid F28B3.  
A:Reference number: Z21469  
A:Accession: T34057  
A:Status: preliminary; translated from GB/EMBL/DDBT  
A:Molecule type: DNA  
A:Residues: 1-918 <CEI>  
A:Cross-references: EMBL:AF003136; PIDN:AB93636.1; GSPDB:GN00019; CESP:F28B3.5  
A:Experimental source: strain Bristol N2; clone F28B3  
C:Genetics:  
A:Gene: CESP:F28B3.5  
A:Map position: 1  
A:introns: 85/3; 129/3; 235/3; 418/2; 482/3; 532/3; 736/3; 829/3; 850/2

Query Match 2.5%; Score 9; DB 2; Length 918;  
Best Local Similarity 100.0%; Pred. No. 1.3;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 211 LDAIVDVT 219  
DB 762 LDAIVDVT 770

RESULT 5  
W8BPG7  
gene 18.7 protein - phage T7  
C:Species: phage T7  
C>Date: 13-Jun-1983 #sequence\_revision 13-Jun-1983 #text\_change 23-Jul-1999  
C:Accession: A04413; S42337  
R:Dunn, J.U.; Thompson, K.  
submitted to the Nucleic Acid Sequence Database, September 1982  
A:Reference number: A94615  
A:Accession: A04413  
A:Molecule type: DNA  
A:Residues: 1-83 <DUN>  
R:Dunn, J.U.; Studier, F.W.  
J. Mol. Biol. 166, 477-535, 1983  
A:Title: Complete nucleotide sequence of bacteriophage T7 DNA and the locations of T7  
A:Reference number: S42283; MUID:83241725  
A:Accession: S42337  
A:Molecule type: DNA  
A:Residues: 1-83 <DUN>  
A:Cross-references: EMBL:V01146; NID:g431187; PIDN:CAA24439.1; PID:g15615  
A:Note: the authors did not translate the codon for residue 1  
C:Genetics:  
A:Gene: 18.7  
A:Map position: 92.73-93.35  
C:Superfamily: phage T7 gene 18.7 protein

Query Match 2.3%; Score 8; DB 1; Length 83;  
Best Local Similarity 100.0%; Pred. No. 1.6;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 305 LSIKKTLP 312  
DB 22 LSIKKTLP 29

RESULT 6  
W8BPT3  
gene 18.7 protein - phage T3  
C:Species: phage T3  
A:Note: host Escherichia coli  
C>Date: 31-Dec-1988 #sequence\_revision 31-Dec-1988 #text\_change 23-Jul-1999  
C:Accession: E23476  
R:Yamada, M.; Fujisawa, H.; Kato, H.; Hamada, K.; Minagawa, T.  
Virology 151, 350-361, 1986  
A:Title: Cloning and sequencing of the genetic right end of bacteriophage T3 DNA.  
A:Reference number: A94339; MUID:86209997  
A:Accession: E23476  
A:Molecule type: DNA  
A:Residues: 1-83 <YAM>  
A:Cross-references: GB:M14784; NID:g215810; PIDN:AA92527.1; PID:g1196765  
R:Yamada, M.; Fujisawa, H.; Kato, H.; Hamada, K.; Minagawa, T.  
Virology 154, 246, 1986  
A:Reference number: A94344  
A:Contents: annotation; erratum; corrections to coding regions  
C:Genetics:  
A:Gene: 18.7  
C:Superfamily: phage T7 gene 18.7 protein

Query Match 2.3%; Score 8; DB 1; Length 83;  
Best Local Similarity 100.0%; Pred. No. 1.6;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 305 LSIKKTLP 312  
DB 22 LSIKKTLP 29

RESULT 7  
H90892  
probable fimbrial chaperone protein precursor [Imported] - Escherichia coli (strain O  
C:Species: Escherichia coli



C:Date: 18-Jul-2001 #sequence\_revision 18-Jul-2001 #text\_change 03-Aug-2001  
 C:Accession: H90892  
 R:Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ichi, K.; Yokoyama, K.; Han, C.G.  
 gesawara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shingawa, H.  
 DNA Res. 8, 11-22, 2001  
 A:Title: Complete genome sequence of enterohemorrhagic *Escherichia coli* O157:H7 and genc  
 A:Reference number: A99629; MUID:21156231; PMID:11258796  
 A:Accession: H90892  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-236 <HAY>  
 A:Cross-references: GB:BA000007, PIDN:BA835535.1, PID:g13361578; GSPDB:GN00154  
 A:Experimental source: strain O157:H7, substrain RMD 0509952  
 C:Genetics:  
 A:Gene: ECS2112  
 C:Superfamily: chaperone protein papp

Query Match 2.3%; Score 8; DB 2; Length 236;  
 Best Local Similarity 100.0%; Pred. No. 4.2;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Oy 193 LTPRIKAT 200  
 |||||||  
 Db 227 LTPRIKAT 234

RESULT 8  
 B87524  
 1-acyl-sn-glycerol-3-phosphate acyltransferase [Imported] - *Caulobacter crescentus*  
 C:Species: *Caulobacter crescentus*  
 C:Date: 20-Apr-2001 #sequence\_revision 20-Apr-2001 #text\_change 20-Apr-2001  
 C:Accession: B87524  
 R:Nierman, W.C.; Feldlyum, T.V.; Paulsen, I.T.; Nelson, K.E.; Eisen, J.; Heideberg, J.  
 B.; Land, M.T.; DeBoy, R.T.; Dodson, R.J.; Durkin, A.S.; Gwin, M.L.; Haft, D.H.; Kolon  
 n, J.; Ermolaeva, M.; White, O.; Salzberg, S.L.; Shapiro, L.; Venter, J.C.; Fraser, C.M.  
 Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001  
 A:Title: Complete genome sequence of *Caulobacter crescentus*.  
 A:Reference number: A87249; MUID:21173698; PMID:11259647  
 A:Accession: B87524  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-236 <STO>  
 A:Cross-references: GB:AE005673; NID:g13423724; PIDN:AAK24190.1; GSPDB:GN00148  
 C:Genetics:  
 A:Gene: CC2219

Query Match 2.3%; Score 8; DB 2; Length 236;  
 Best Local Similarity 100.0%; Pred. No. 4.2;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Oy 158 VIFPECTR 165  
 |||||||  
 Db 146 VIFPECTR 153

RESULT 9  
 A85725  
 Probable fimbrial chaperone protein z2201 [Imported] - *Escherichia coli* (strain O157:H7,  
 C:Species: *Escherichia coli*  
 C:Date: 16-Feb-2001 #sequence\_revision 16-Feb-2001 #text\_change 14-Sep-2001  
 C:Accession: A85725  
 R:Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhew  
 Miller, L.; Grothbeck, E.J.; Davis, N.W.; Lm, A.; Dinalanta, E.; Potamousis, K.; Apodaca,  
 Nature 409, 529-533, 2001  
 A:Title: Genome sequence of enterohemorrhagic *Escherichia coli* O157:H7.  
 A:Reference number: A85480; MUID:21074935; PMID:11206551  
 A:Accession: A85725  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-239 <STO>  
 A:Cross-references: GB:AE005174; NID:g12515164; PIDN:AAG56261.1; GSPDB:GN00145; OMGP:222

A:Experimental source: strain O157:H7, substrain EDL933  
 C:Genetics:  
 A:Gene: Z2201  
 C:Superfamily: chaperone protein papp

Query Match 2.3%; Score 8; DB 2; Length 239;  
 Best Local Similarity 100.0%; Pred. No. 4.3;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Oy 193 LTPRIKAT 200  
 |||||||  
 Db 230 LTPRIKAT 237

RESULT 10  
 A71478  
 Probable metal dependent hydrolase - *Chlamydia trachomatis* (serotype D, strain UW3/Cx  
 C:Species: *Chlamydia trachomatis*  
 C:Date: 13-Sep-1998 #sequence\_revision 13-Sep-1998 #text\_change 08-Oct-1999  
 C:Accession: A71478  
 R:Stephens, R.S.; Kalman, S.; Lammel, C.J.; Fan, J.; Marathe, R.; Aravind, L.; Mitche  
 Science 282, 754-759, 1998  
 A:Title: Genome sequence of an obligate intracellular pathogen of humans: *Chlamydia t*  
 A:Reference number: A71570; MUID:99000809  
 A:Accession: A71478  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-262 <ARN>  
 A:Cross-references: GB:AE001344; GB:AE001273; NID:g3329188; PIDN:AAC68333.1; PID:g332  
 A:Experimental source: serotype D, strain UW-3/Cx  
 C:Genetics:  
 A:Gene: YycJ

Query Match 2.3%; Score 8; DB 2; Length 262;  
 Best Local Similarity 100.0%; Pred. No. 4.6;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Oy 191 HVLTPRIK 198  
 |||||||  
 Db 202 HVLTPRIK 209

RESULT 11  
 G71334  
 Probable lysophosphatidic acid acyltransferase - *Syphilis spirochete*  
 C:Species: *Treponema pallidum* subsp. *pallidum* (*Syphilis spirochete*)  
 C:Date: 24-Jul-1998 #sequence\_revision 24-Jul-1998 #text\_change 05-Nov-1999  
 C:Accession: G71334  
 R:Fraser, C.M.; Norris, S.J.; Weinstock, G.M.; White, O.; Sutton, G.G.; Dodson, R.; G  
 rson, J.; Khalak, H.; Richardson, D.; Howell, J.K.; Chidambaram, M.; Utterback, T.; M  
 they, L.; Weidman, J.; Smith, H.O.; Venter, J.C.  
 Science 281, 375-388, 1998  
 A:Title: Complete genome sequence of *Treponema pallidum*, the *Syphilis spirochete*.  
 A:Reference number: A71250; MUID:98332770  
 A:Accession: G71334  
 A:Status: preliminary  
 A:Molecule type: nucleic acid sequence not shown; translation not shown  
 A:Molecule type: DNA  
 A:Residues: 1-282 <COL>  
 A:Cross-references: GB:AE001215; GB:AE000520; NID:g3322631; PIDN:AAC65346.1; PID:g332  
 A:Experimental source: strain Nichols  
 C:Genetics:  
 A:Gene: TP0361

Query Match 2.3%; Score 8; DB 2; Length 282;  
 Best Local Similarity 100.0%; Pred. No. 5;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Oy 158 VIFPECTR 165  
 |||||||  
 Db 170 VIFPECTR 177

## RESULT 12

S54641

probable membrane protein YDR018c - yeast (*Saccharomyces cerevisiae*)

N/Alternate names: hypothetical protein D3246; hypothetical protein P2F396; hypothetical

C/Species: *Saccharomyces cerevisiae*

C/Date: 08-Jul-1995 #sequence\_revision 01-Sep-1995 #text\_change 20-Jun-2000

C/Accession: S54641; S63425; S67831; S72116

R/Edman, K.; Brown, D.; Hamlyn, N.; Bowman, S.

submitted to the EMBL Data Library, May 1995

A/Reference number: S54638

A/Accession: S54641

A/Molecule type: DNA

A/Residues: 1-396 &lt;DED&gt;

A/Cross-references: EMBL:249770; NID:g840867; PIDN:CAA9843.1; PID:g840871

A/Experimental source: strain AB972

R/Elde, L.G.; Sander, C.; Prydz, H.

submitted to the EMBL Data Library, February 1996

A/Description: Sequencing and analysis of a 35.4 kb region on the left arm of chromosome

A/Reference number: S63416

A/Accession: S63425

A/Molecule type: DNA

A/Residues: 1-396 &lt;EID&gt;

A/Cross-references: EMBL:X95966; NID:g1216215; PIDN:CAA65210.1; PID:g1216225

R/Prydz, H.; Elde, L.G.

submitted to the Protein Sequence Database, July 1996

A/Reference number: S67822

A/Accession: S67831

A/Molecule type: DNA

A/Residues: 1-396 &lt;PRY&gt;

A/Cross-references: EMBL:274314; NID:g1431443; PIDN:CAA9838.1; PID:g1431444; MIPS:YDR01

A/Experimental source: strain S288C

R/Elde, L.G.; Sander, C.; Prydz, H.

Yeast 12, 1085-1090, 1996

A/Title: Sequencing and analysis of a 35.4 kb region on the left arm of chromosome IV fr

A/Reference number: S72107; MUID:97051598

A/Accession: S72116

A/Status: nucleic acid sequence not shown; translation not shown

A/Molecule type: DNA

A/Residues: 1-396 &lt;EIM&gt;

A/Cross-references: EMBL:X95966; NID:g1216215; PIDN:CAA65210.1; PID:g1216225

A/Note: The nucleotide sequence was submitted to the EMBL Data Library, February 1996

C/Genetics:

A/Map position: 4R

A/Note: YDR018c

C/Superfamily: probable membrane protein YBR042c

C/Keywords: transmembrane protein

F:27-43/Domain: transmembrane #status predicted &lt;TM1&gt;

F:69-85/Domain: transmembrane #status predicted &lt;TM2&gt;

F:376-392/Domain: transmembrane #status predicted &lt;TM3&gt;

DB 265 LDAIYDVT 272

OY 211 LDAIYDVT 218

## RESULT 13

C87586

metal ion efflux membrane fusion protein family [imported] - *Caulobacter crescentus*C/Species: *Caulobacter crescentus*

C/Date: 20-Apr-2001 #sequence\_revision 20-Apr-2001 #text\_change 30-Jun-2001

C/Accession: C87586

R/Merman, W.C.; Feldblum, T.V.; Paulsen, I.T.; Nelson, K.E.; Eisen, J.; Heidelberg, J.

B.; Laub, M.T.; Deboy, R.T.; Dodson, R.J.; Durkin, A.S.; Gilm, M.L.; Haft, D.H.; Kolton

n, J.; Ermolaeva, M.; White, O.; Salzberg, S.L.; Shapiro, L.; Venter, J.C.; Fraser, C.M.

Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001

A/Title: Complete Genome Sequence of *Caulobacter crescentus*.

A/Reference number: A87249; MUID:21173698; PMID:11259647

A/Accession: C87586

A/Status: preliminary

A/Molecule type: DNA

A/Residues: 1-411 &lt;STO&gt;

A/Cross-references: GB:AE005673; NID:g13424311; PIDN:AAK24667.1; GSPDB:GN00148

C/Genetics:

C/Superfamily: nickel-cobalt resistance determinant structural protein CnrB; 1lipoyl/b

Query Match 2.3%; Score 8; DB 2; Length 411;

Best Local Similarity 100.0%; Pred. No. 7;

Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 316 ILSGTAG 323  
DB 380 ILSGTAG 387

## RESULT 14

T49607

guanine deaminase (Gda) related protein [imported] - *Neurospora crassa*

N/Alternate names: protein B3E4.200

C/Species: *Neurospora crassa*

C/Date: 02-Jun-2000 #sequence\_revision 02-Jun-2000 #text\_change 02-Jun-2000

C/Accession: T49607

R/Schulte, U.; Algn, V.; Hohelsel, J.; Brandt, P.; Fartmann, B.; Holland, R.; Nyakatu

submitted to the Protein Sequence Database, May 2000

A/Reference number: Z25022

A/Accession: T49607

A/Status: preliminary

A/Molecule type: DNA

A/Residues: 1-424 &lt;SCH&gt;

A/Cross-references: EMBL:AL355931; GSPDB:GN00116; NCSP:B3E4.200

A/Experimental source: BAC clone B3E4; strain OR74A

C/Genetics:

A/Map position: 6

Query Match 2.3%; Score 8; DB 2; Length 424;

Best Local Similarity 100.0%; Pred. No. 7.2;

Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 189 LKHVTPR 196  
DB 178 LKHVTPR 185

## RESULT 15

RNA methyltransferase BH0687 [imported] - *Bacillus halodurans* (strain C-125)C/Species: *Bacillus halodurans*

C/Date: 01-Dec-2000 #sequence\_revision 01-Dec-2000 #text\_change 15-Jun-2001

C/Accession: G83735

R/Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Masui, N.; Fujii, F.; H

Nucleic Acids Res. 28, 4317-4331, 2000

A/Title: Complete genome sequence of the alkaliphilic bacterium *Bacillus halodurans* a

A/Reference number: A83650; MUID:20512582; PMID:11058132

A/Accession: G83735

A/Status: preliminary

A/Molecule type: DNA

A/Residues: 1-458 &lt;STO&gt;

A/Cross-references: GB:AP001509; GB:BA000004; NID:g10173176; PIDN:BA04406.1; GSPDB:G

A/Experimental source: strain C-125

C/Genetics:

A/Map position:

Query Match 2.3%; Score 8; DB 2; Length 458;

Best Local Similarity 100.0%; Pred. No. 7.7;

Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 167 NPEQTKVL 174  
| | | | |  
Db 291 NPEQTKVL 298

Search completed: August 28, 2002, 11:20:08  
Job time: 261 sec

**THIS PAGE BLANK (USPTO)**

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 28, 2002, 11:09:02 ; Search time 75.33 Seconds  
(without alignments)  
520.497 Million cell updates/sec

Title: US-09-853-526-4  
Perfect score: 353  
Sequence: 1 MRYLPSVLLGTAPFYVLA.....YVNWIVGTLGLMTWTKA 353

Scoring table:  
Gapop 60.0 , Gapext 60.0

Searched: 747574 seqs, 111073796 residues

Word size : 0

Total number of hits satisfying chosen parameters: 747574

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database :

A.Geneseq\_032802:\*  
1: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1980.DAT:\*  
2: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1981.DAT:\*  
3: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1982.DAT:\*  
4: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1983.DAT:\*  
5: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1984.DAT:\*  
6: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1985.DAT:\*  
7: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1986.DAT:\*  
8: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1987.DAT:\*  
9: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1988.DAT:\*  
10: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1989.DAT:\*  
11: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1990.DAT:\*  
12: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1991.DAT:\*  
13: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1992.DAT:\*  
14: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1993.DAT:\*  
15: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1994.DAT:\*  
16: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1995.DAT:\*  
17: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1996.DAT:\*  
18: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1997.DAT:\*  
19: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1998.DAT:\*  
20: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1999.DAT:\*  
21: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA2000.DAT:\*  
22: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA2001.DAT:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	353	100.0	353	22	AAB85532
2	353	100.0	353	22	AAB85552
3	353	100.0	364	20	AAV36729
4	353	100.0	364	22	AAV50128
5	353	100.0	364	22	AAU00665
6	352	99.7	352	20	AAV36728
7	290	82.2	353	22	AAV36591
8	269	76.2	269	22	AAE09653
9	252	71.4	353	22	AAV36595
10	252	71.4	372	22	AAV41377
11	230	65.2	291	20	AAV36744

12	228	64.6	228	20	AAV36740
13	199	56.4	261	20	AAV36745
14	185	52.4	185	20	AAV36753
15	185	52.4	315	20	AAV36751
16	184	52.1	300	20	AAV36752
17	154	43.6	182	20	AAV36750
18	116	32.9	238	20	AAV36743
19	85	24.1	96	20	AAV36749
20	70	19.8	257	22	ABG01360
21	70	19.8	980	22	ABG11977
22	62	17.6	66	20	AAV36748
23	62	17.6	68	20	AAV36747
24	62	17.6	77	20	AAV36742
25	62	17.6	90	20	AAV36746
26	53	15.0	97	22	ABG12868
27	53	15.0	1104	22	ABG12107
28	32	9.1	354	20	AAV36741
29	9	2.5	9	20	AAV36734
30	9	2.5	157	22	AAV50130
31	9	2.5	237	18	AAV55540
32	9	2.5	240	19	AAV98477
33	9	2.5	242	18	AAV55441
34	8	2.3	9	20	AAV36737
35	8	2.3	239	22	ABB52710
36	8	2.3	396	21	AAV99484
37	8	2.3	557	22	ABB71470
38	8	2.3	1820	22	ABB62370
39	7	2.0	12	16	AAR87705
40	7	2.0	12	20	AAV43149
41	7	2.0	65	16	AAR87720
42	7	2.0	82	16	AAV43164
43	7	2.0	82	16	AAR87718
44	7	2.0	82	16	AAR87719
45	7	2.0	82	20	AAV43160

#### ALIGNMENTS

RESULT 1	
AAB85532	standard; protein: 353 AA.
XX	
AC	AAB85532;
XX	
DT	25-SEP-2001 (first entry)
XX	
DE	Human secreted protein (clone id HLTH084) .
XX	
XX	
KW	Secreted protein; immunosuppressive; antiarthritic; antirheumatic;
KW	antiproliferative; cytostatic; cardiant; vasotrophic; cerebroprotective;
KW	neotropic; neuroprotective; antibacterial; virucide; fungicide; human;
KW	ophthalmological; gene therapy.
XX	
OS	Homo sapiens.
XX	
PN	WO200155430-A1.
XX	
PD	02-AUG-2001.
XX	
PF	17-JAN-2001; 2001WO-US01431.
XX	
PR	31-JAN-2000; 2000US-0179065.
PR	04-FEB-2000; 2000US-0180628.
PR	12-SEP-2000; 2000US-0231968.
PA	(HUMA-) HUMAN GENOME SCI INC.
XX	
PI	Rosen CA, Komatsoulis GH, Baker KP, Birse CE, Soppet DR, Olsen HS;
PI	Moore PA, Wei P, Edner R, Duan DR, Shi Y, Choi GH, Fiscella M;
XX	Ni J, Ruben SM, Barash SC;
DR	WPI; 2001-476220/51.

Truncated PGI prot  
PGI splice variant  
PGI splice variant  
PGI splice variant  
PGI splice variant  
PGI splice variant  
PGI splice variant  
PGI splice variant  
PGI splice variant  
Novel human diagno  
Novel human diagno  
Novel human diagno  
PGI splice variant  
PGI splice variant  
PGI splice variant  
Novel human diagno  
Novel human diagno  
Mouse PGI protein  
PGI box 3 motif.  
Lysophosphatidic a  
H. pylori ORF 12ge  
H. pylori GHPO 903  
H. pylori ORF 06gp  
Acyl glycerol tran  
Escherichia coli P  
Yeast acyltransfer  
Drosophila melanog  
Coconut AG3PPT try  
Coconut 1-acylglyc  
Coconut LPAAT clon  
1-acylglycerol-3-p  
Coconut LPAAT clon  
Coconut LPAAT clon  
1-acylglycerol-3-p

DR N-PSDB: AAH46942.  
XX  
PT 17 isolated nucleic acid molecules encoding human secreted proteins,  
XX used to preventing, treating or ameliorating a medical condition -  
XX  
PS Claim 11: Page 450-451: 482pp: English.  
XX  
XX The invention provides novel human secreted proteins and polynucleotides  
CC encoding them. The secreted proteins can be expressed by standard  
CC recombinant methodology. The secreted proteins and polynucleotides are  
CC used to prevent, treat or ameliorate a medical condition in e.g. humans,  
CC mice, rabbits, goats, horses, cats, dogs, chickens or sheep. They can  
CC also be used in diagnosing a pathological condition. The antibodies to  
CC the proteins can also be used in alleviating symptoms associated with the  
CC disorders and in diagnostic immunoassays e.g. radioimmunoassays or enzyme  
CC linked immunoassay assays (ELISA). Disorders which are diagnosed or  
CC treated include autoimmune diseases e.g. rheumatoid arthritis,  
CC hyperproliferative disorders e.g. neoplasms of the breast or liver,  
CC cardiovascular disorders e.g. cardiac arrest, cerebrovascular disorders  
CC e.g. cerebral ischemia, angiodenesis, nervous system disorders e.g.  
CC Alzheimer's disease, infections caused by bacteria, viruses and fungi and  
CC ocular disorders e.g. corneal infection. The polypeptides can also be  
CC used to aid wound healing and epithelial cell proliferation, to prevent  
CC skin aging due to sunburn, to maintain organs before transplantation, for  
CC supporting cell culture of primary tissues, to regenerate tissues and in  
CC chemotaxis. The polypeptides can also be used as a food additive or  
CC preservative to increase or decrease storage capabilities. The present  
CC sequence represents a human secreted protein.  
XX  
SQ Sequence 353 AA:  
  
Query Match 100.0%; Score 353; DB 22; Length 353;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 353; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 1 MRYLPSVNLGTPYVLAAGWVRLSAFLPAFYQALDRLCYQSWVLEFFENYTG 60  
DB 1 mrylpsvnlgtapyvlagwvrlsalfparfyqaldrlcyqswvleffenyttg 60  
QY 61 VOILYGDLPKRNENIYLANHOSTVDMIVADILAIROMALGHVRYVLEKGLMPLXYGC 120  
DB 61 vqlllygdldpknkenillylanhgstvdwivadilaistqnalghvryvlekgkwlplxygc 120  
QY 121 YFAOHGSIYVKRSKAKFNEKEMRNKLOSYVDAGTPMYLVIFPEGTRYNPEQTKVLSASQAF 180  
DB 121 yfaohgsgiyvkrsakfnekemrnklqsyvdagtpmylvifpegrtrynpeqtkvlsasqaf 180  
QY 181 AAOGLAVLKHVLTLPRIKATHVAFDCKKNYLDIAYDTVYVEGDDGQORRESPTMTFEL 240  
DB 181 aagglavlkhlvtlprlkathvafdcmknyldaiydtvvyegkddggqrresptmtefl 240  
QY 241 CKECPKTHIHIDRIKDVPEOEHRMRMLHERPEIKDKMLIEFYESPDPERRKRFPGKS 300  
DB 241 ckecpkthihidridkdvpeoehrmrmlherpeikdkmliefyespdperrkrfpgks 300  
QY 301 VNSKLSIKTKLPMSMLISGLTAGMLMTDAGRKLVTNTWYIGTLLGCLMWTIKA 353  
DB 301 vnskslisktklpmsmlisgltagmlmtdagrklyvntwiygllgclmwltika 353  
  
RESULT 2  
AAB85352  
ID AAB85552 standard; protein; 353 AA.  
XX  
AC AAB85552:  
XX  
DT 25-SEP-2001 (first entry)  
XX  
XX Human secreted protein (clone id HSLIA81).  
XX  
KM Secreted protein; immunosuppressive; antiarthritic; antirheumatic;  
KW antiproliferative; cyostatic; cardiac; vasotropic; cerebroprotective;

KW neurotropic; neuroprotective; antibacterial; virucide; fungicide; human;  
KW ophthalmological; gene therapy.  
XX  
OS Homo sapiens.  
XX  
PN W020015430-A1.  
XX  
PD 02-AUG-2001.  
XX  
PF 17-JAN-2001; 2001MO-US01431.  
XX  
XX 31-JAN-2000; 2000US-0179065.  
PR 04-FEB-2000; 2000US-0180628.  
PR 12-SEP-2000; 2000US-0231968.  
XX  
PA (HUMA-) HUMAN GENOME SCI INC.  
XX  
PI Rosen CA, Komatsu G, Baker KP, Birse CE, Soppet DR, Olsen HS;  
PI Moore PA, Wei P, Eber R, Duan DR, Shi Y, Choi GH, Fiscella M;  
PI Ni J, Ruben SM, Barash SC;  
XX  
XX WPI: 2001-476220/51.  
DR N-PSDB: AAH46942.  
XX  
XX  
PT 17 isolated nucleic acid molecules encoding human secreted proteins,  
PT used to preventing, treating or ameliorating a medical condition -  
XX  
PS Claim 11: Page 465-466; 482pp: English.  
XX  
XX The invention provides novel human secreted proteins and polynucleotides  
CC encoding them. The secreted proteins can be expressed by standard  
CC recombinant methodology. The secreted proteins and polynucleotides are  
CC used to prevent, treat or ameliorate a medical condition in e.g. humans,  
CC mice, rabbits, goats, horses, cats, dogs, chickens or sheep. They can  
CC also be used in diagnosing a pathological condition. The antibodies to  
CC the proteins can also be used in alleviating symptoms associated with the  
CC disorders and in diagnostic immunoassays e.g. radioimmunoassays or enzyme  
CC linked immunoassay assays (ELISA). Disorders which are diagnosed or  
CC treated include autoimmune diseases e.g. Rheumatoid arthritis,  
CC hyperproliferative disorders e.g. neoplasms of the breast or liver,  
CC cardiovascular disorders e.g. cardiac arrest, cerebrovascular disorders  
CC e.g. cerebral ischemia, angiodenesis, nervous system disorders e.g.  
CC Alzheimer's disease, infections caused by bacteria, viruses and fungi and  
CC ocular disorders e.g. corneal infection. The polypeptides can also be  
CC used to aid wound healing and epithelial cell proliferation, to prevent  
CC skin aging due to sunburn, to maintain organs before transplantation, for  
CC supporting cell culture of primary tissues, to regenerate tissues and in  
CC chemotaxis. The polypeptides can also be used as a food additive or  
CC preservative to increase or decrease storage capabilities. The present  
CC sequence represents a human secreted protein.  
XX  
SQ Sequence 353 AA:  
  
Query Match 100.0%; Score 353; DB 22; Length 353;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 353; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 1 MRYLPSVNLGTPYVLAAGWVRLSAFLPAFYQALDRLCYQSWVLEFFENYTG 60  
DB 1 mrylpsvnlgtapyvlagwvrlsalfparfyqaldrlcyqswvleffenyttg 60  
QY 61 VOILYGDLPKRNENIYLANHOSTVDMIVADILAIROMALGHVRYVLEKGLMPLXYGC 120  
DB 61 vqlllygdldpknkenillylanhgstvdwivadilaistqnalghvryvlekgkwlplxygc 120  
QY 121 YFAOHGSIYVKRSKAKFNEKEMRNKLOSYVDAGTPMYLVIFPEGTRYNPEQTKVLSASQAF 180  
DB 121 yfaohgsgiyvkrsakfnekemrnklqsyvdagtpmylvifpegrtrynpeqtkvlsasqaf 180  
QY 181 AAOGLAVLKHVLTLPRIKATHVAFDCKKNYLDIAYDTVYVEGDDGQORRESPTMTFEL 240  
DB 181 aagglavlkhlvtlprlkathvafdcmknyldaiydtvvyegkddggqrresptmtefl 240

OY	241	CKEPEKIHINHRIDRDKDVEDEQHNMRRLHEREREIKODKLIEFESPPDERKKRPGRKS	300
Db	241	CKEPKIhhhdrrdkkdvpeeghmrtwtherfelkdmliefesppderkkirpgrks	300
OY	301	VNSKLSIKKTLPMSLILSGLTAGMLMTDAGRKLIVVNTWINGTLLGLAMWTIRA	353
Db	301	vnskslksktklpslmllisgltagmmtcdagrkilyntwtlygtlllgclwtlrka	353
<hr/>			
RESULT	3		
ID	AAY36729		
XX	AAV36729 standard; Protein; 364 AA.		
AC	AAV36729;		
XX			
DJ	27-SEP-1999 (first entry)		
XX			
DE	Human PGI protein sequence.		
KW	PGI gene; biallelic marker; human; PSA; PGI-related biallelic marker;		
XX	cancer; prostate cancer; diagnosis; therapy; prostate specific antigen.		
OS	Homo sapiens.		
XX			
PN	MO9932644-AA.		
PD			
XX	01-JUL-1999.		
PF			
PR	22-DEC-1998; 98WO-IB02133.		
XX			
PR	09-SEP-1998; 98US-0099658.		
XX	22-DEC-1997; 97US-0996306.		
PA	(GEST ) GENSET.		
XX			
P1	Blumenfeld M, Bougueleret L, Chumakov I, Cohen D;		
DR	WPJ: 1999-405178/34.		
N-PSDB:	AAZ00870.		
PT			
XX			
XX	Use of a prostate cancer associated gene and biallelic markers		
XX	derived from it		
PS	Claim 7; Page 190-191; 385pp; English.		
CC	The invention relates to a mammalian PGI gene and protein, and a set of		
CC	PGI biallelic markers. The PGI polynucleotide and biallelic markers are		
CC	used in a hybridization assay, a sequencing assay, or in an		
CC	allele-specific amplification assay for determining the identity of a		
CC	nucleotide at a PGI-related biallelic marker. The methods can be used to		
CC	detect and to assess the risk of developing cancer or prostate cancer.		
CC	Early-stage diagnosis of prostate cancer relies on prostate specific		
CC	antigen (PSA) dosage. However, the effectiveness of this is limited due		
CC	to its inability to discriminate between malignant and non-malignant		
CC	affectons of the organ. A need exists for both a reliable diagnostic		
CC	procedure which would enable early-stage diagnosis, and for preventative		
CC	and curative treatments of the disease. The PGI gene can be used for		
CC	detection of prostate cancer, and the risk of developing it in the		
CC	future, and can also be used to determine therapies for the disease.		
SQ	Sequence 364 AA;		
<hr/>			
Query Match	100.0%;	Score 353;	DB 20; Length 364;
Best Local Similarity	100.0%;	Pred. No. 0;	
Matches 353;	Conservative 0;	Mismatches 0;	Indels 0; Gaps
OY	1	MRYLPSVVLGTPTVTVIAMGVWRLLSAPFAFYALDDRLCYQVSATLEFFENYTG	60
Db	12	mrylpsvvlvgtcaptyvlaqgwvrllsatlfparfyaldorrlcyqysmvllftfnytg	71
OY	61	VOILLYGDLPPKNRNIIYYLANHSTVDWIVADLIAIQNALGHVRVYLAKGLWPLPYGC	120

Db	72	vqlllygdipknkenillylaangstcvwiyadilairqnalghvrylkegikwlpipygc	131
Oy	121	YFAHGIGITVYRSKRFKEKEKRNKLQSYVAGTPEMVLVIPEEGTRVNEOTKVLASQAF	180
Db	132	yfaqdgqylgvyraskfknekemrnklqsyvadqgipmylvifpeqtrvnypeqtkvlasaqaf	191
Oy	181	AAQGLAVLKHVLTPEPRIKATHVAFDCKKNVLDALIDYTVVYEGKDDGGQRRSPMTMTEFL	240
Db	192	aagrglavlkhvltprikathvafcdcknuyldalydvtcvyegkddggqrrspmtmtefl	251
Oy	241	CKEECKIHIIHIDRIDKKDVPEQEHMRKWLHERREIKDKMLIERYESPDPERKKRPFGKS	300
Db	252	ckeockihihidridkdkdvpeeqehmrwvherfeikdkmliefyespdperkrifpgks	311
Oy	301	VNSKLSIKTKLPMSLISGLFAGMLMTDAGKTLVNWIMYIGTLGLCLMTVRIKA	353
Db	312	vnsklsiktklpmsllisglfagmlmtdagrklyvntwiytllgclwvrika	364
RESULT 4			
AA50128	ID	AA50128 standard; Protein; 364 AA.	
XX	AC	AA50128;	
XX	DT	21-DEC-2001 (first entry)	
XX	DE	Human acyltransferase 27417.	
XX	XX		
KW	ACyltransferase 27417; human; cell proliferation; cell migration;		
KW	cell differentiation; ovary cancer; brain cancer; colon cancer;		
KW	lung cancer; tumour; metastasis; sarcoma; carcinoma;		
KW	adenocarcinoma; antitumour; lipid metabolism; diagnosis; therapy.		
XX	XX		
OS		Homo sapiens.	
XX	XX		
FT	Key	Location/Qualifiers	
FT	Peptide	1..49	
FT	Protein	/label= Signal_peptide	
FT	Domain	/label= Mature_protein	
FT	Domain	/label= N-terminal_domain	
FT	Domain	/label= Transmembrane_domain	
FT	Domain	/label= C-terminal_domain	
FT	Domain	/label= Acyltransferase_domain	
FT	Modified-site	/note= "Asn is N-glycosylated"	
FT	Modified-site	/note= "O-phosphorylated by cAMP- and cGMP-dependent protein kinase"	
FT	Modified-site	11..13	
FT	Modified-site	/note= "O-phosphorylated by protein kinase C"	
FT	Modified-site	144..146	
FT	Modified-site	/note= "O-phosphorylated by protein kinase C"	
FT	Modified-site	205..207	
FT	Modified-site	/note= "O-phosphorylated by protein kinase C"	
FT	Modified-site	317..319	
FT	Modified-site	/note= "O-phosphorylated by protein kinase C"	
FT	Modified-site	361..363	
FT	Modified-site	/note= "O-phosphorylated by protein kinase C"	
FT	Modified-site	95..98	
FT	Modified-site	/note= "O-phosphorylated by casein kinase II"	
FT	Modified-site	158..161	
FT	Modified-site	/note= "O-phosphorylated by casein kinase II"	
FT	Modified-site	246..249	
FT	Modified-site	/note= "O-phosphorylated by casein kinase II"	
FT	Modified-site	82..89	
FT	Modified-site	/note= "O-phosphorylated by tyrosine kinase"	

Db	72	vqlllygdipknkenillylaangstcvwiyadilairqnalghvrylkegikwlpipygc	131
Oy	121	YFAHGIGITVYRSKRFKEKEKRNKLQSYVAGTPEMVLVIPEEGTRVNEOTKVLASQAF	180
Db	132	yfaqdgqylgvyraskfknekemrnklqsyvadqgipmylvifpeqtrvnypeqtkvlasaqaf	191
Oy	181	AAQGLAVLKHVLTPEPRIKATHVAFDCKKNVLDALIDYTVVYEGKDDGGQRRSPMTMTEFL	240
Db	192	aagrglavlkhvltprikathvafcdcknuyldalydvtcvyegkddggqrrspmtmtefl	251
Oy	241	CKEECKIHIIHIDRIDKKDVPEQEHMRKWLHERREIKDKMLIERYESPDPERKKRPFGKS	300
Db	252	ckeockihihidridkdkdvpeeqehmrwvherfeikdkmliefyespdperkrifpgks	311
Oy	301	VNSKLSIKTKLPMSLISGLFAGMLMTDAGKTLVNWIMYIGTLGLCLMTVRIKA	353
Db	312	vnsklsiktklpmsllisglfagmlmtdagrklyvntwiytllgclwvrika	364
RESULT 4			
AA50128	ID	AA50128 standard; Protein; 364 AA.	
XX	AC	AA50128;	
XX	DT	21-DEC-2001 (first entry)	
XX	DE	Human acyltransferase 27417.	
XX	XX		
KW	ACyltransferase 27417; human; cell proliferation; cell migration;		
KW	cell differentiation; ovary cancer; brain cancer; colon cancer;		
KW	lung cancer; tumour; metastasis; sarcoma; carcinoma;		
KW	adenocarcinoma; antitumour; lipid metabolism; diagnosis; therapy.		
XX	XX		
OS		Homo sapiens.	
XX	XX		
FT	Key	Location/Qualifiers	
FT	Peptide	1..49	
FT	Protein	/label= Signal_peptide	
FT	Domain	/label= Mature_protein	
FT	Domain	/label= N-terminal_domain	
FT	Domain	/label= Transmembrane_domain	
FT	Domain	/label= C-terminal_domain	
FT	Domain	/label= Acyltransferase_domain	
FT	Modified-site	/note= "Asn is N-glycosylated"	
FT	Modified-site	/note= "O-phosphorylated by cAMP- and cGMP-dependent protein kinase"	
FT	Modified-site	11..13	
FT	Modified-site	/note= "O-phosphorylated by protein kinase C"	
FT	Modified-site	144..146	
FT	Modified-site	/note= "O-phosphorylated by protein kinase C"	
FT	Modified-site	205..207	
FT	Modified-site	/note= "O-phosphorylated by protein kinase C"	
FT	Modified-site	317..319	
FT	Modified-site	/note= "O-phosphorylated by protein kinase C"	
FT	Modified-site	361..363	
FT	Modified-site	/note= "O-phosphorylated by protein kinase C"	
FT	Modified-site	95..98	
FT	Modified-site	/note= "O-phosphorylated by casein kinase II"	
FT	Modified-site	158..161	
FT	Modified-site	/note= "O-phosphorylated by casein kinase II"	
FT	Modified-site	246..249	
FT	Modified-site	/note= "O-phosphorylated by casein kinase II"	
FT	Modified-site	82..89	
FT	Modified-site	/note= "O-phosphorylated by tyrosine kinase"	

Query Match	Best Local Similarity	Score 353:	DB 22:	Length 364:
Matches 353:	Conservative 0:	Mismatches 0:	Indels 0:	Gaps 0:
1	MRYLLPVSVLGTAFYTVLAWGVMFLSAFLPARFYQALDRLYCYVQSMVLFPEENYTG 60	100.0%:	Score 353:	DB 22:
12	mcylilpvsvlilgtapcytlaagvwlilafiparfylgaldtdllycvysmvlffennytg 71	100.0%:	Score 353:	DB 22:

OY	61	VQILLTGLPEFNKRNKITTYLANHSTVDWYIADOLAIRONALGHVRYVLEBGLAKWPIRYGC	120
OY	72	VQILLTGLPEFNKRNKITTYLANHSTVDWYIADOLAIRONALGHVRYVLEBGLAKWPIRYGC	131
OY	121	YFAOHGIGYVYRSKAFNPEKEMRNKLQSYVAGPMTLVIFPEECTRYNPEOTVYLSASQAF	180
DB	132	YFAOHGIGYVYRSKAFNPEKEMRNKLQSYVAGPMTLVIFPEECTRYNPEOTVYLSASQAF	191
OY	181	AAQGLAVLKHVLTPTRIKATHVAFDCMKNYLAIYDVYVYEGSKDDGGQRESPTMTEFL	240
DB	192	AAQGLAVLKHVLTPTRIKATHVAFDCMKNYLAIYDVYVYEGSKDDGGQRESPTMTEFL	251
OY	241	CKECPKLIHIDRIDKKDVPDEEHEHMRWMLHEEFELKDKMLLEFYESPPEERKKPPGKS	300
DB	252	CKECPKLIHIDRIDKKDVPDEEHEHMRWMLHEEFELKDKMLLEFYESPPEERKKPPGKS	311
OY	301	VNSKLSIKTKLPSMLILSGLTAGMLMTDAGRKLYVTWVITWITGLGLAWTIKA	353
DB	312	VNSKLSIKTKLPSMLILSGLTAGMLMTDAGRKLYVTWVITWITGLGLAWTIKA	364
RESULT	5		
AAU00665			
AAU00665		standard; Protein; 364 AA.	
AAU00665:			
DT	07-SEP-2001	(first entry)	
DE		Human lysophosphatidic acid acyltransferase isoform LPAAT-epsilon.	
KW		Lysophosphatidic acid acyltransferase epsilon; LPAAT-epsilon; LPA; PA;	
KM		lysophosphatidic acid; phosphatidic acid; acylation; cellular activation;	
KW		phospholipid signaling; mitogenesis; inflammation; autoimmune disease;	
KM		oncology; cancer; obesity; gene therapy.	
OS		Homo sapiens.	
XX			
XX			
PN		MO200134782-A1.	
PD		17-MAY-2001.	
XX			
PF		02-NOV-2000; 2000MO-US03193.	
XX			
PR		09-NOV-1999; 99US-0436919.	
XX			
XX		(CELL-) CELL THERAPEUTICS INC.	
XX			
PI		Leung DW;	
XX			
DR		WPI: 2001-335920/35.	
DR		N-PSDB: AAS00649.	
XX			
PT		Novel isolated human isoform of lysophosphatidic acid	
PT		acyltransferase-epsilon useful for diagnostic, therapeutic and	
XX		screening purposes	
PS		Claim 1; Fig 1; 48pp; English.	
XX			
CC		The sequence represents a human lysophosphatidic acid acyltransferase	
CC		(LPAAT) isoform, LPAAT-epsilon. LPAAT catalyses the acylation of	
CC		lysophosphatidic acid (LPA) to phosphatidic acid (PA). LPA and PA have	
CC		been identified as phospholipid signaling molecules that affect a wide	
CC		range of biological responses. PA is involved in cellular activation and	
CC		mitogenesis. Compounds that block PA generation and hence diminish lipid	
CC		biosynthesis and the signal involved in cell activation are of	
CC		therapeutic interest in the areas of inflammation and oncology (e.g.	
CC		autoimmune diseases and cancer) as well as obesity treatment.	
CC		LPAAT-epsilon and its corresponding DNA can be used in screening assays	
CC		to detect agents that stimulate or inhibit the activity of LPAAT and,	
CC		therefore, PA. The DNA is useful in tests to detect the presence or	
CC		expression of LPAAT-epsilon in relation to certain diseases and	



CC conditions, and in disease prevention and treatment. The sequences of the  
CC invention are also useful for diagnosis of diseases and conditions in  
CC which the expression of LPAAT enzyme is abnormal.

XX Sequence 364 AA;

Query Match 100.0%; Score 353; DB 22; Length 364;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 353; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 MRYLLPSVLLGTAPTYVLAMGWRLLSAFLPAFYQALDRLCYQSWLFEFFENYTG 60  
DB 12 mryllpsvllgtaptyvlamgwrrllsalfparfyqalddrlcyqswmlfffenytg 71  
OY 61 VOILLYGDLPKKNENITYLANHOSYVDWIVADILAIRONALGHVRYLKEGLKWLPLYGC 120  
DB 72 vqilllygdlpkknennyilanhsyvdwivadilaairgnalghvrylkeglkwlplygc 131  
OY 121 YFAOHGIYVKRSKAFKNEKEMRNKLSQSYDAGTPMYLVIPPEGTRNYPEQTKVLSAQAF 180  
DB 132 yfaohgiyvkrsakfnekemrnlksqsydagtpmylvippegtrnypegtkvlssaqaaf 191  
OY 181 AAOQGLAVLKHVLTLPRIKATHVAFDCMKNYLDIYDVTVVYEGKDGQGRRESPTMTFEL 240  
DB 192 aaqrglavlkhvltprikathvafdcmknyldiaydtvvyeqkdqggrresptmtefl 251  
OY 241 CKCEPKIHIDRIDKDDVPEOEHRMRWLHERPEIKDKMLIEFYESPDERRKRFRPGKS 300  
DB 252 ckcepkihidridkddvpeegehmrwlherfeikdkmliefyespderrkrfrpgks 311  
OY 301 VNSKLSIKKTLPMSMLISGLTAGMLMTDAGRKLTVNWTWYIGTLIGCLMTWITKA 353  
DB 312 vnsksikktlpmsmlisgltagmlmtdagrklvntwtwyigtligclwtwika 364

## RESULT 6

AA36728  
ID AAY36728 standard; Protein; 352 AA.

XX AC AAY36728;

XX DT 27-SEP-1999 (first entry)

XX DE Human PGI protein sequence.

KW PGI gene; biallelic marker; human; PSA; PGI-related biallelic marker;  
KW cancer; prostate cancer; diagnosis; therapy; prostate specific antigen.

XX OS Homo sapiens.

XX PW WO9932644-A2.

XX PD 01-JUL-1999.

XX PF 22-DEC-1998; 98MO-IB02133.

XX PR 09-SEP-1998; 98US-0099658.

XX PR 22-DEC-1997; 97US-0996306.

XX PA (GEST ) GENSET.

XX PI Blumenfeld M, Bougueleret L, Chumakov I, Cohen D;

XX DR WPI; 1999-405178/34.

XX DR N-PSDB; AA200872.

XX PT Use of a prostate cancer associated gene and biallelic markers

XX PS derived from it

XX PS Claim 7; Page 190; 385pp; English.

CC The invention relates to a mammalian PGI gene and protein, and a set of

CC PGI biallelic markers. The PGI polynucleotide and biallelic markers are  
CC used in a hybridisation assay, a sequencing assay, or in an  
CC allele-specific amplification assay for determining the identity of a  
CC nucleotide at a PGI-related biallelic marker. The methods can be used to  
CC detect and to assess the risk of developing cancer or prostate cancer.  
CC Early-stage diagnosis of prostate cancer relies on prostate specific  
CC antigen (PSA) dosage. However, the effectiveness of this is limited due  
CC to its inability to discriminate between malignant and non-malignant  
CC affections of the organ. A need exists for both a reliable diagnostic  
CC procedure which would enable early-stage diagnosis, and for preventative  
CC and curative treatments of the disease. The PGI gene can be used for  
CC detection of prostate cancer, and the risk of developing it in the  
CC future, and can also be used to determine therapies for the disease.

XX Sequence 352 AA;

Query Match 99.7%; Score 352; DB 20; Length 352;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 352; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 MRYLLPSVLLGTAPTYVLAMGWRLLSAFLPAFYQALDRLCYQSWLFEFFENYTG 60  
DB 1 mryllpsvllgtaptyvlamgwrrllsalfparfyqalddrlcyqswmlfffenytg 60  
OY 61 VOILLYGDLPKKNENITYLANHOSYVDWIVADILAIRONALGHVRYLKEGLKWLPLYGC 120  
DB 61 vqilllygdlpkknennyilanhsyvdwivadilaairgnalghvrylkeglkwlplygc 120  
OY 121 YFAOHGIYVKRSKAFKNEKEMRNKLSQSYDAGTPMYLVIPPEGTRNYPEQTKVLSAQAF 180  
DB 121 yfaohgiyvkrsakfnekemrnlksqsydagtpmylvippegtrnypegtkvlssaqaaf 180  
OY 181 AAOQGLAVLKHVLTLPRIKATHVAFDCMKNYLDIYDVTVVYEGKDGQGRRESPTMTFEL 240  
DB 181 aaqrglavlkhvltprikathvafdcmknyldiaydtvvyeqkdqggrresptmtefl 240  
OY 241 CKCEPKIHIDRIDKDDVPEOEHRMRWLHERPEIKDKMLIEFYESPDERRKRFRPGKS 300  
DB 241 ckcepkihidridkddvpeegehmrwlherfeikdkmliefyespderrkrfrpgks 300  
OY 301 VNSKLSIKKTLPMSMLISGLTAGMLMTDAGRKLTVNWTWYIGTLIGCLMTWITKA 352  
DB 301 vnsksikktlpmsmlisgltagmlmtdagrklvntwtwyigtligclwtwika 352

## RESULT 7

AAM39591  
ID AAM39591 standard; Protein; 353 AA.

XX AC AAM39591;

XX DT 22-OCT-2001 (first entry)

XX DE Human polypeptide SEQ ID NO 2736.

XX KW

XX KW

XX KW

XX KW

XX OS Homo sapiens.

XX PN WO200153312-A1.

XX PD 26-JUL-2001.

XX PF 26-DEC-2000; 2000MO-US34263.

XX PR 21-JAN-2000; 2000US-0488725.

XX PR 25-APR-2000; 2000US-0552317.

PR 09-JUL-2000; 2000US-0598042.  
PR 19-JUL-2000; 2000US-0620312.  
PR 03-AUG-2000; 2000US-0653450.  
PR 14-SEP-2000; 2000US-0662191.  
PR 19-OCT-2000; 2000US-0693036.  
PR 29-NOV-2000; 2000US-0727344.  
XX  
XX (HYSE-) HYSEQ INC.  
XX  
XX Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;  
PI Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J;  
PI Zhao QA, Zhou P, Goodrich R, Drmanac RT;  
XX  
XX WPI: 2001-442253/47.  
DR N-PSDB: AA158747.  
XX  
XX Novel nucleic acids and polypeptides, useful for treating disorders  
PT such as central nervous system injuries -  
XX  
XX Example 4: SEQ ID NO 2736; 10078bp; English.  
XX  
XX The invention relates to human nucleic acids (AA157798-AA161369) and  
CC the encoded polypeptides (AA158642-AA162213) with nootropic.  
CC Immunosuppressant and cyostatic activity. The polynucleotides are useful  
CC in gene therapy. A composition containing a polypeptide or polynucleotide  
CC of the invention may be used to treat diseases of the peripheral nervous  
CC system, such as peripheral nervous injuries, peripheral neuropathy and  
CC localised neuropathies and central nervous system diseases, such as  
CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic  
CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the  
CC utilisation of the activities such as: Immune system suppression,  
CC Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic  
CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,  
CC assays for receptor activity, arthritis and inflammation, leukaemias and  
CC C.N.S disorders.  
CC Note: The sequence data for this patent did not form part of the printed  
CC specification.  
XX  
XX Sequence 353 AA;  
SQ

Query Match 82.2%; Score 290; DB 22; Length 353;  
Best Local Similarity 100.0%; Pred. No. 3.8e-287;  
Matches 290; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 MRYLLPSVVLGGTAPTVLAWGVRLSAFLPARFYQALDRLCYVQSMYLPFFENYTG 60  
DB 1 mryllpsvvlggtaptyvlawgvrrllsafiparfqalddrlcyvqsmlylffenytg 60  
QY 61 VOILLYDLPKKNENIYYLANHSTVDWIVADILAIKONALGHVRYLKEGLKMLPLXGC 120  
DB 61 vqilllygdllprknenillylanhgstvdwivadilairqnalghvrylkeglkwlplygc 120  
QY 121 YFNOHGITYVRSKFNKEKEMRNKLOSYVDAGTPMYLYIPEEGRYNPEQKKVYSASQAF 180  
DB 121 yfnohgityvrsakfnkekemrnlklsyvdagtrpmylylpeegtrypnpeqtkvlsasqaf 180  
QY 181 AAORGLVNLKHLVLPRIKATHVAFDCMKNYLDALYDVTYVVEGKDDGQRRSESTMTTEFL 240  
DB 181 aagrglavnlkhlvlprikathvafdcmknyldaldydvtyvveggddgqrrsestlmtefl 240  
QY 241 CKECPRKHIDRIKDKVPEDEOHMRRLHEFEIKDKMLIEFYESPDP 290  
DB 241 ckecpkhlhldridkdkvpeegohmrmlhefeikdkmliefyespdp 290

RESULT 8  
AAE09653  
ID AAE09653 standard; Protein: 269 AA.  
XX  
AC AAE09653;  
XX  
DT 19-NOV-2001 (first entry)

XX  
DE Human gene 15 encoded lipid metabolism protein RHASB32, SEQ ID NO:47.  
XX  
XX  
KW cerebroprotective; hepatotropic; neuroprotective; cardiatic;  
KW immune disorder; autoimmune disease; rheumatoid arthritis; glossitis;  
KW systemic lupus erythematosus; conjunctivitis; inflammatory disorder;  
KW respiratory disorder; asthma; allergy; CNS disorder; Alzheimer's disease;  
KW Parkinson's disease; atherosclerosis; cardiovascular disorder; cancer;  
KW coronary disease; familial hypercholesterolaemia; hyperlipidaemia;  
KW hemapoleptic disorder; hypolipidaemia; lipodosis; Gaucher's disease;  
KW Tay-sach's disease; mental retardation; gene therapy; antisense therapy.  
XX  
XX Homo sapiens.  
OS  
XX  
XX WO20015203-A1.  
PN  
XX  
PD 02-AUG-2001.  
XX  
XX 17-JAN-2001; 2001WO-US01327.  
PE  
XX  
XX 31-JAN-2000; 2000US-0179065.  
PR 04-FEB-2000; 2000US-0180628.  
PR 24-FEB-2000; 2000US-0184664.  
PR 02-MAR-2000; 2000US-0186350.  
PR 16-MAR-2000; 2000US-0189874.  
PR 18-MAR-2000; 2000US-0190076.  
PR 18-APR-2000; 2000US-0198123.  
PR 19-MAY-2000; 2000US-0205515.  
PR 07-JUN-2000; 2000US-0209467.  
PR 28-JUN-2000; 2000US-0214886.  
PR 30-JUN-2000; 2000US-0215135.  
PR 07-JUL-2000; 2000US-0216647.  
PR 07-JUL-2000; 2000US-0216880.  
PR 11-JUL-2000; 2000US-0217487.  
PR 11-JUL-2000; 2000US-0217496.  
PR 14-JUL-2000; 2000US-0218290.  
PR 26-JUL-2000; 2000US-0220963.  
PR 26-JUL-2000; 2000US-0220964.  
PR 14-AUG-2000; 2000US-0224518.  
PR 14-AUG-2000; 2000US-0224519.  
PR 14-AUG-2000; 2000US-0225213.  
PR 14-AUG-2000; 2000US-0225214.  
PR 14-AUG-2000; 2000US-0225215.  
PR 14-AUG-2000; 2000US-0225266.  
PR 14-AUG-2000; 2000US-0225267.  
PR 14-AUG-2000; 2000US-0225268.  
PR 14-AUG-2000; 2000US-0225270.  
PR 14-AUG-2000; 2000US-0225447.  
PR 14-AUG-2000; 2000US-0225757.  
PR 14-AUG-2000; 2000US-0225758.  
PR 14-AUG-2000; 2000US-0225759.  
PR 18-AUG-2000; 2000US-0226279.  
PR 22-AUG-2000; 2000US-0226681.  
PR 22-AUG-2000; 2000US-0226688.  
PR 22-AUG-2000; 2000US-0227182.  
PR 22-AUG-2000; 2000US-0227009.  
PR 23-AUG-2000; 2000US-0228924.  
PR 30-AUG-2000; 2000US-0229287.  
PR 01-SEP-2000; 2000US-0229343.  
PR 01-SEP-2000; 2000US-0229344.  
PR 01-SEP-2000; 2000US-0229345.  
PR 05-SEP-2000; 2000US-0229509.  
PR 05-SEP-2000; 2000US-0229513.  
PR 06-SEP-2000; 2000US-0230437.  
PR 06-SEP-2000; 2000US-0230438.  
PR 08-SEP-2000; 2000US-0231242.  
PR 08-SEP-2000; 2000US-0231243.  
PR 08-SEP-2000; 2000US-0231244.  
PR 08-SEP-2000; 2000US-0231413.  
PR 08-SEP-2000; 2000US-0231414.  
PR 08-SEP-2000; 2000US-0232080.  
PR 08-SEP-2000; 2000US-0232081.  
PR 12-SEP-2000; 2000US-0231968.

PR 14-SEP-2000; 2000US-0233397.  
 PR 14-SEP-2000; 2000US-0233398.  
 PR 14-SEP-2000; 2000US-0233399.  
 PR 14-SEP-2000; 2000US-0233400.  
 PR 14-SEP-2000; 2000US-0233401.  
 PR 14-SEP-2000; 2000US-0233063.  
 PR 14-SEP-2000; 2000US-0233064.  
 PR 14-SEP-2000; 2000US-0233065.  
 PR 21-SEP-2000; 2000US-0234223.  
 PR 21-SEP-2000; 2000US-0234274.  
 PR 25-SEP-2000; 2000US-0234997.  
 PR 25-SEP-2000; 2000US-0234998.  
 PR 26-SEP-2000; 2000US-0235484.  
 PR 27-SEP-2000; 2000US-0235834.  
 PR 27-SEP-2000; 2000US-0235836.  
 PR 29-SEP-2000; 2000US-0236327.  
 PR 29-SEP-2000; 2000US-0236367.  
 PR 29-SEP-2000; 2000US-0236368.  
 PR 29-SEP-2000; 2000US-0236369.  
 PR 29-SEP-2000; 2000US-0236370.  
 PR 02-OCT-2000; 2000US-0236802.  
 PR 02-OCT-2000; 2000US-0237037.  
 PR 02-OCT-2000; 2000US-0237038.  
 PR 02-OCT-2000; 2000US-0237039.  
 PR 02-OCT-2000; 2000US-0237040.  
 PR 13-OCT-2000; 2000US-0239935.  
 PR 13-OCT-2000; 2000US-0239937.  
 PR 20-OCT-2000; 2000US-0240960.  
 PR 20-OCT-2000; 2000US-0241221.  
 PR 20-OCT-2000; 2000US-0241785.  
 PR 20-OCT-2000; 2000US-0241786.  
 PR 20-OCT-2000; 2000US-0241787.  
 PR 20-OCT-2000; 2000US-0241808.  
 PR 20-OCT-2000; 2000US-0241809.  
 PR 20-OCT-2000; 2000US-0241826.  
 PR 01-NOV-2000; 2000US-0244617.  
 PR 01-NOV-2000; 2000US-0244617.  
 PR 08-NOV-2000; 2000US-0246474.  
 PR 08-NOV-2000; 2000US-0246475.  
 PR 08-NOV-2000; 2000US-0246476.  
 PR 08-NOV-2000; 2000US-0246477.  
 PR 08-NOV-2000; 2000US-0246478.  
 PR 08-NOV-2000; 2000US-0246523.  
 PR 08-NOV-2000; 2000US-0246524.  
 PR 08-NOV-2000; 2000US-0246525.  
 PR 08-NOV-2000; 2000US-0246526.  
 PR 08-NOV-2000; 2000US-0246527.  
 PR 08-NOV-2000; 2000US-0246528.  
 PR 08-NOV-2000; 2000US-0246532.  
 PR 08-NOV-2000; 2000US-0246609.  
 PR 08-NOV-2000; 2000US-0246610.  
 PR 08-NOV-2000; 2000US-0246611.  
 PR 08-NOV-2000; 2000US-0246613.  
 PR 17-NOV-2000; 2000US-0249207.  
 PR 17-NOV-2000; 2000US-0249208.  
 PR 17-NOV-2000; 2000US-0249209.  
 PR 17-NOV-2000; 2000US-0249210.  
 PR 17-NOV-2000; 2000US-0249211.  
 PR 17-NOV-2000; 2000US-0249212.  
 PR 17-NOV-2000; 2000US-0249213.  
 PR 17-NOV-2000; 2000US-0249214.  
 PR 17-NOV-2000; 2000US-0249215.  
 PR 17-NOV-2000; 2000US-0249216.  
 PR 17-NOV-2000; 2000US-0249217.  
 PR 17-NOV-2000; 2000US-0249218.  
 PR 17-NOV-2000; 2000US-0249244.  
 PR 17-NOV-2000; 2000US-0249245.  
 PR 17-NOV-2000; 2000US-0249264.  
 PR 17-NOV-2000; 2000US-0249265.  
 PR 17-NOV-2000; 2000US-0249297.  
 PR 17-NOV-2000; 2000US-0249299.  
 PR 17-NOV-2000; 2000US-0249300.  
 PR 01-DEC-2000; 2000US-0250160.  
 PR 01-DEC-2000; 2000US-0250391.

PR 05-DEC-2000; 2000US-0251030.  
 PR 05-DEC-2000; 2000US-0251988.  
 PR 05-DEC-2000; 2000US-0251989.  
 PR 06-DEC-2000; 2000US-0251479.  
 PR 06-DEC-2000; 2000US-0251856.  
 PR 08-DEC-2000; 2000US-0251856.  
 PR 08-DEC-2000; 2000US-0251869.  
 PR 08-DEC-2000; 2000US-0251869.  
 PR 08-DEC-2000; 2000US-0251989.  
 PR 11-DEC-2000; 2000US-0251990.  
 PR 05-JAN-2001; 2001US-0259678.  
 XX  
 XX (HUMA-) HUMAN GENOME SCI INC.  
 PA  
 PI Rosen CA, Barash SC, Ruben SM.  
 XX  
 XX WPI; 2001-457116/49.  
 DR N-PSDB; AAD16620.  
 XX  
 XX Isolated lipid metabolism polypeptide for screening to identify  
 PT antagonists and agonists that may enhance or block activities mediated  
 PT by lipid metabolism proteins and also for testing and detection e.g.  
 PT diagnosis -  
 XX  
 XX Claim 11; SEQ ID No 47; 453bp; English.  
 PS  
 XX  
 CC AAD16606-AAD16627 represent cDNAs corresponding to human lipid metabolism  
 CC protein genes, and AAD09639-AAD09660 represent the proteins they encode.  
 CC AAD16628-AAD16647 represent human lipid metabolism genomic DNAs. The  
 CC genes and their corresponding proteins are useful for preventing,  
 CC treating or ameliorating medical conditions such as immune disorders,  
 CC inflammatory disorders (e.g., conjunctivitis, glossitis), autoimmune  
 CC diseases (e.g., rheumatoid arthritis, systemic lupus erythematosus),  
 CC respiratory disorders (asthma, allergy), CNS disorders (e.g., Alzheimer's  
 CC disease, Parkinson's disease), haematopoietic disorders, cardiovascular  
 CC disorders (e.g., atherosclerosis, coronary disease), hyperlipidaemias,  
 CC hypolipidaemia, lipidoses, familial hypercholesterolaemia, Gaucher's  
 CC disease, Tay-sachs's disease, diabetes mellitus, cancer and other enzyme  
 CC abnormalities. They are used for the prevention of mental retardation  
 CC and the enlargement of spleen and liver. They are used in gene therapy  
 CC and antisense therapy. The present sequence represents a human lipid  
 CC metabolism protein of the invention.  
 CC Note: The sequence data for this patent did not form part of the printed  
 CC specification, but was obtained in electronic format directly from WIPO  
 CC at ftp.wipo.int/pub/published\_pct\_sequences.  
 XX  
 XX Sequence 269 AA;  
 SQ  
 Query Match 76.2%; Score 269; DB 22; Length 269;  
 Best Local Similarity 100.0%; Pred. No. 7.7e-266;  
 Matches 269; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 85 TYDVIADILAIRONALGHVRYVLKGLKWLPLGYCYFAOHGSIYVRSARFNEKEMRNK 144  
 Db 1 tvdviadvadilairgnalghvryvlkeglkwlplgygyfaahgsllyvrsakfknekmrnk 60  
 QY 145 LOSYVAGTPMYLYVIRPEEGRIYRNEBOTKYLSAQARAQGLAVLKHVLPRTKATHAVAR 204  
 Db 61 lgsyvdaqtpmylyvifpgegrlyrnpetkylsasqafaagylavlvhlprikathavaf 120  
 QY 205 DCKKNYLDAIYDVTVYEGGDGGRRESPTWEPICKECPKIHIIHIDRDKDVEEOE 264  
 Db 121 dckknlydaivdvtyvgygdggrresptweflickecpklnhldridkdvpeee 180  
 QY 265 HMRRWLHERFEIKDKMLIEFESPDDERRKRFPKGSVNSKLSIKTKLPMSLISGLTAGM 324  
 Db 181 hmrrwlherfelkdkmliefespddeerrkrfpgksvnskslksltpmsllisgltagm 240  
 QY 325 LMTDAGRKLVTWYITGLTGLCWVITKA 353  
 Db 241 lmtdagrklyvntwlytglglcwvltika 269

RESULT 9  
AAB93595  
ID AAB93595 standard; Protein: 353 AA.  
XX  
AC AAB93595;  
XX  
DT 26-JUN-2001 (first entry)  
XX  
DE Human protein sequence SEQ ID NO:13028.  
XX  
KM Human; primer; detection; diagnosis; antisense therapy; gene therapy.  
XX  
OS Homo sapiens.  
XX  
PN EP1074617-A2.  
XX  
PD 07-FEB-2001.  
XX  
PE 28-JUL-2000; 2000EP-0116126.  
XX  
PF 29-JUL-1999; 99JP-0248036.  
XX  
PR 27-AUG-1999; 99JP-0300253.  
XX  
PR 11-JAN-2000; 2000JP-0118776.  
XX  
PR 02-MAY-2000; 2000JP-0183767.  
XX  
PR 09-JUN-2000; 2000JP-0241899.  
XX  
PA (HELI-) HELIX RES INST.  
XX  
PI Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;  
PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;  
DR WPI: 2001-318749/34.  
XX  
PT Primer sets for synthesizing polynucleotides, particularly the 5602  
PT full-length cDNAs defined in the specification, and for the detection  
PT and/or diagnosis of the abnormality of the proteins encoded by the  
PT full-length cDNAs -  
XX  
PS Claim 8; SEQ ID 13028; 2537pp + CD ROM; English.  
XX  
XX The present invention describes primer sets for synthesizing 5602  
CC full-length cDNAs defined in the specification. Where a primer set  
CC comprises: (a) an oligo-dT primer and an oligonucleotide complementary  
CC to the complementary strand of a polynucleotide which comprises one of  
CC the 5602 nucleotide sequences defined in the specification, where the  
CC oligonucleotide comprises at least 15 nucleotides; or (b) a combination  
CC of an oligonucleotide comprising a sequence complementary to the  
CC complementary strand of a polynucleotide which comprises a 5'-end  
CC sequence and an oligonucleotide comprising a sequence complementary to a  
CC polynucleotide which comprises a 3'-end sequence, where the  
CC oligonucleotide comprises at least 15 nucleotides and the combination of  
CC the 5'-end sequence/3'-end sequence is selected from those defined in  
CC the specification. The primer sets can be used in antisense therapy and  
CC in gene therapy. The primers are useful for synthesizing polynucleotides,  
CC particularly full-length cDNAs. The primers are also useful for the  
CC detection and/or diagnosis of the abnormality of the proteins encoded by  
CC the full-length cDNAs. The primers allow obtaining of the full-length  
CC cDNAs easily without any specialised methods. AAH03166 to AAH13628 and  
CC AAH13633 to AAH18742 represent human cDNA sequences; AAB92446 to  
CC AAB95893 represent human amino acid sequences; and AAH13629 to AAH13632  
CC represent oligonucleotides, all of which are used in the exemplification  
CC of the present invention.  
XX  
SQ Sequence 353 AA;

Query Match 71.4%; Score 252; DB 22; Length 353;  
Best Local Similarity 99.7%; Pred. No. 2.2e-248;  
Matches 352; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 1 MRLLPSVVLGTPPYVLAAGWRLSAPLPARFYALDRLXCYQSWLFFFEWYTG 60  
XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX

Db 1 mryllpsvvl1gtapcyvlawgvvrllsafiiparfyalddrlycyvqswvlfffenytg 60  
QY 61 VOILLGYDLPKKNENIYYLANHSTVDMIVADILAIRQNALGHVRYVLEKGLKMLPLYGC 120  
XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX  
Db 61 vqilllygd1pknkeni1y1lanhsgvtvawdlairqnalghvryv1lkeglkw1plygc 120  
QY 121 YFAOHGGIYYKRSKAKFENEKRNKLSQSYVDAAGTPMYLVIFPEGTRRYPEOTKVLASQAF 180  
XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX  
Db 121 yfaohggilyvkrskafenekrnklsqsyvdagtpmylvifpegrtrypedtkvlasaqaf 180  
QY 181 AAORGGLVLRKHLVLPRIKATHVAFDCKKNYLDAIDYDTVYVEGDKDGGORRESPTMEFL 240  
XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX  
Db 181 aagrgglavlkrhlvlprikatvhvafdccknnyldaidydvtyvvegkdggrresptmefl 240  
QY 241 CKECPKIHIDRIKRDVPEDEEHMRRLHERFEIKDKMLIEFESPDERRKRFPKGS 300  
XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX  
Db 241 ckecpkih1hidr1dkrdvpeede1hmr1lherfe1kdkml1efespderrkr1fpgks 300  
QY 301 VNSKLSIKKTLPSMLISGLTAGMLMTDAGRKLYVNTWITGTLICLWVTIKA 353  
XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX  
Db 301 vnskslksk1t1psml1sg1tagmlmtcdagrklvntw1tgytlgclwv1ika 353  
RESULT 10  
AAM41377  
ID AAM41377 standard; Protein: 372 AA.  
XX  
AC AAM41377;  
XX  
XX  
XX  
DT 22-OCT-2001 (first entry)  
XX  
DE Human polypeptide SEQ ID NO 6308.  
XX  
KM Human; nootropic; immunosuppressant; cytostatic; gene therapy; cancer;  
XX peripheral nervous system; neuropathy; central nervous system; CNS;  
XX Alzheimer's; Parkinson's disease; Huntington's disease; hemostatic;  
XX amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;  
XX chemokine; thrombolytic; drug screening; arthritis; inflammation;  
XX leukaemia.  
XX  
XX Homo sapiens.  
XX  
XX  
XX WO200153312-A1.  
XX  
XX 26-JUL-2001.  
XX  
XX 26-DEC-2000; 2000MO-US34263.  
XX  
XX 21-JAN-2000; 2000US-0488725.  
XX  
XX 25-APR-2000; 2000US-052317.  
XX  
XX 09-JUL-2000; 2000US-0598042.  
XX  
XX 19-JUL-2000; 2000US-0620312.  
XX  
XX 03-AUG-2000; 2000US-0653450.  
XX  
XX 14-SEP-2000; 2000US-0662191.  
XX  
XX 19-OCT-2000; 2000US-0693036.  
XX  
XX 29-NOV-2000; 2000US-0727344.  
XX  
XX (HYSE-) HYSEQ INC.  
XX  
XX Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;  
XX Wang JT, Wang Z, Weinman T, Xu C, Xue AJ, Yang Y, Zhang J;  
XX Zhao QA, Zhou P, Goodrich R, Drmanac RT;  
XX  
XX N-PSDB; AA160533.  
XX  
XX WPI: 2001-442253/47.  
XX  
XX Novel nucleic acids and polypeptides, useful for treating disorders  
XX such as central nervous system injuries -  
XX  
XX Example 2; SEQ ID NO 6308; 10078pp; English.  
XX  
XX The invention relates to human nucleic acids (AA157798-AA161369) and  
XX the encoded polypeptides (AAM38642-AAM42213) with nootropic,

CC immunosuppressant and cytostatic activity. The polynucleotides are useful  
 CC in gene therapy. A composition containing a polypeptide or polynucleotide  
 CC of the invention may be used to treat diseases of the peripheral nervous  
 CC system, such as peripheral nervous injuries, peripheral neuropathy and  
 CC localised neuropathies and central nervous system diseases, such as  
 CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic  
 CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the  
 CC utilisation of the activities such as: Immune system suppression,  
 CC activation/inhibition activity, chemotactic/chemokinetic activity, haemostatic  
 CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,  
 CC assays for receptor activity, arthritis and inflammation, leukemias and  
 CC C.N.S disorders.  
 CC Note: The sequence data for this patent did not form part of the printed  
 CC specification.

XX Sequence 372 AA;

Query Match 71.4%; Score 252; DB 22; Length 372;  
 Best Local Similarity 99.7%; Pred. No. 2.3e-248;  
 Matches 352; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MRLPLPSVVLGTRPTVYVLAAGWVRLLSARLPARFYALDDRLCYVQSNVLEFFENYTG 60  
 |||||||  
 DB 20 MYLLPSVVLGTRPTVYVLAAGWVRLLSARLPARFYALDDRLCYVQSNVLEFFENYTG 79  
 |||||||  
 QY 61 VOILLYGDLPKKNENIYYLANHOSTVDIVADILAIRONALGHVRYLKEGKMLPLYGC 120  
 |||||||  
 DB 80 VQLLLYGDLPKKNENIYYLANHOSTVDIVADILAIRONALGHVRYLKEGKMLPLYGW 139  
 |||||||  
 QY 121 YFAOHGIVYKRSKAFNEKEEMKRNKLSQSYVDAGTPMVLVIEPEGRVPEQTKVLSASQAF 180  
 |||||||  
 DB 140 YFAGHGIVYKRSKAFNEKEEMKRNKLSQSYVDAGTPMVLVIEPEGRVPEQTKVLSASQAF 199  
 |||||||  
 QY 181 AAGRGVLAHLVLPRTKATHVAFDCMKNYLDATYDVTVVYEGKDDGQRRSEPTMTFEL 240  
 |||||||  
 DB 200 AAGRGVLAHLVLPRTKATHVAFDCMKNYLDATYDVTVVYEGKDDGQRRSEPTMTFEL 259  
 |||||||  
 QY 241 CKECPKIHIDRIDKDVPEEGEEMRRWLHERFEIKDKMLIEFESPDPERRRKRFPGKS 300  
 |||||||  
 DB 260 CKECPKIHIDRIDKDVPEEGEEMRRWLHERFEIKDKMLIEFESPDPERRRKRFPGKS 319  
 |||||||  
 QY 301 VNSKLSIKTKLPMSLLISGLTAGMLMTDAGRKLIVNNTWIVYGTLLGCLMTWTKA 353  
 |||||||  
 DB 320 VNSKLSIKTKLPMSLLISGLTAGMLMTDAGRKLIVNNTWIVYGTLLGCLMTWTKA 372

RESULT 11

AAV36744

ID AAV36744 standard; Protein; 291 AA.

XX AAV36744;

DT 27-SEP-1999 (first entry)

DE PGI splice variant allele protein sequence.

KW PGI gene; biallelic marker; human; PSA; PGI-related biallelic marker;  
 KW cancer; prostate cancer; diagnosis; therapy; prostate specific antigen.

OS Homo sapiens.

XX MO9932644-A2.

PN 01-JUL-1999.

PF 22-DEC-1998; 98WO-IB02133.

PR 09-SEP-1998; 98US-0099658.  
 PR 22-DEC-1997; 97US-0996306.

XX (GEST ) GENSET.

XX

PI Blumenfeld M, Bougueleret L, Chumakov I, Cohen D;  
 XX WPI: 1999-405178/34.  
 DR N-PSDB: AA200970.

PT Use of a prostate cancer associated gene and biallelic markers  
 PT derived from it

PS Claim 7; Page 250-251; 385pp; English.

CC The invention relates to a mammalian PGI gene and protein, and a set of  
 CC PGI biallelic markers. The PGI polynucleotide and biallelic markers are  
 CC used in a hybridisation assay, a sequencing assay, or in an  
 CC allele-specific amplification assay for determining the identity of a  
 CC nucleotide at a PGI-related biallelic marker. The methods can be used to  
 CC detect and to assess the risk of developing cancer or prostate cancer.  
 CC Early-stage diagnosis of prostate cancer relies on prostate specific  
 CC antigen (PSA) dosage. However, the effectiveness of this is limited due  
 CC to its inability to discriminate between malignant and non-malignant  
 CC affections of the organ. A need exists for both a reliable diagnostic  
 CC procedure which would enable early-stage diagnosis, and for preventative  
 CC and curative treatments of the disease. The PGI gene can be used for  
 CC detection of prostate cancer, and the risk of developing it in the  
 CC future, and can also be used to determine therapies for the disease.

XX Sequence 291 AA;

Query Match 65.2%; Score 230; DB 20; Length 291;  
 Best Local Similarity 100.0%; Pred. No. 5e-226;  
 Matches 230; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 124 OHGIVYKRSKAFNEKEEMKRNKLSQSYVDAGTPMVLVIEPEGRVPEQTKVLSASQAFQAQ 183  
 |||||||  
 DB 62 QHGIVYKRSKAFNEKEEMKRNKLSQSYVDAGTPMVLVIEPEGRVPEQTKVLSASQAFQAQ 121  
 |||||||  
 QY 184 RGLVLAHLVLPRTKATHVAFDCMKNYLDATYDVTVVYEGKDDGQRRSEPTMTFELCKE 243  
 |||||||  
 DB 122 RGLVLAHLVLPRTKATHVAFDCMKNYLDATYDVTVVYEGKDDGQRRSEPTMTFELCKE 181  
 |||||||  
 QY 244 CPKIHIDRIDKDVPEEGEEMRRWLHERFEIKDKMLIEFESPDPERRRKRFPGKSVNS 303  
 |||||||  
 DB 182 CPKIHIDRIDKDVPEEGEEMRRWLHERFEIKDKMLIEFESPDPERRRKRFPGKSVNS 241  
 |||||||  
 QY 304 KLSIKTKLPMSLLISGLTAGMLMTDAGRKLIVNNTWIVYGTLLGCLMTWTKA 353  
 |||||||  
 DB 242 KLSIKTKLPMSLLISGLTAGMLMTDAGRKLIVNNTWIVYGTLLGCLMTWTKA 291

RESULT 12

AAV36740

ID AAV36740 standard; Protein; 228 AA.

XX AAV36740;

DT 27-SEP-1999 (first entry)

DE Truncated PGI protein sequence.

KW PGI gene; biallelic marker; human; PSA; PGI-related biallelic marker;  
 KW cancer; prostate cancer; diagnosis; therapy; prostate specific antigen.

OS Homo sapiens.

XX MO9932644-A2.

PN 01-JUL-1999.

PF 22-DEC-1998; 98WO-IB02133.

PR 09-SEP-1998; 98US-0099658.  
 PR 22-DEC-1997; 97US-0996306.

XX

PA (GEST ) GENSET.  
 XX Blumenfeld M, Bougueleret L, Chumakov I, Cohen D;  
 XX WPI: 1999-405178/34.  
 DR N-PSDB: AAZ00926.  
 XX  
 PT Use of a prostate cancer associated gene and biallelic markers  
 PT derived from it  
 XX  
 PS Claim 7; Page 208; 385pp; English.  
 XX  
 CC The invention relates to a mammalian PGI gene and protein, and a set of  
 CC PGI biallelic markers. The PGI polynucleotide and biallelic markers are  
 CC used in a hybridisation assay, a sequencing assay, or in an  
 CC allele-specific amplification assay for determining the identity of a  
 CC nucleotide at a PGI-related biallelic marker. The methods can be used to  
 CC detect and to assess the risk of developing cancer or prostate cancer.  
 CC Early-stage diagnosis of prostate cancer relies on prostate specific  
 CC antigen (PSA) dosage. However, the effectiveness of this is limited due  
 CC to its inability to discriminate between malignant and non-malignant  
 CC affections of the organ. A need exists for both a reliable diagnostic  
 CC procedure which would enable early-stage diagnosis, and for preventative  
 CC and curative treatments of the disease. The PGI gene can be used for  
 CC detection of prostate cancer, and the risk of developing it in the  
 CC future, and can also be used to determine therapies for the disease.  
 CC  
 XX  
 SQ Sequence 228 AA;

Query Match 64.6%; Score 228; DB 20; Length 228;  
 Best Local Similarity 100.0%; Pred. No. 4.4e-224;  
 Matches 228; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 MRYLLPSVVLGTAPTYVLANGVWRLLSAFLPARFYQALDDRLCYVQSMFLPEENTG 60  
 Db 1 mryllpsvvlgtaptvylangvwrllsaflparfyqalddrlcyvqsmvllfpeentg 60  
 OY 61 VOILLYGDLPKKNENIYLANHOSTVDMIVADILAIROMALGHVRYVKEGLKMLPLXGC 120  
 Db 61 vqilllygdlpkknenillylanhostvdmivadilaironalghvryvkeglkmlplxgc 120  
 OY 121 YFAOHGGIYVRSKAFNKEKRNKLOSYVDAGTPLYLVIPEEGTRRYNPEQTKVLSASQAF 180  
 Db 121 yfahggilyvrsakfnkekrnklosyvda gtpmlylvifpeegtrrynpegtkvlisaqaf 180  
 OY 181 AAOGLAVLKHVLPRIKATHVAFDCMKNYDAIYDVTVVVEGDDG 228  
 Db 181 aagglavlkhlvprikathvafdcmknydaitydvltvvvegddg 228

RESULT 13  
 AAY36745  
 ID AAY36745 standard; Protein: 261 AA.  
 XX  
 AC AAY36745;  
 XX  
 DT 27-SEP-1999 (first entry)  
 XX  
 DE PGI splice variant allele protein sequence.  
 XX  
 KW PGI gene: biallelic marker; human; PSA; PGI-related biallelic marker;  
 KW cancer; prostate cancer; diagnosis; therapy; prostate specific antigen.  
 OS Homo sapiens.  
 XX  
 PN WO9932644-A2.  
 XX  
 PD 01-JUL-1999.  
 XX  
 PF 22-DEC-1998; 98WO-IB02133.  
 XX  
 PR 09-SEP-1998; 98US-0099658.

PR 22-DEC-1997; 97US-0996306.  
 XX  
 PA (GEST ) GENSET.  
 XX Blumenfeld M, Bougueleret L, Chumakov I, Cohen D;  
 XX WPI: 1999-405178/34.  
 DR N-PSDB: AAZ00971.  
 XX  
 PT Use of a prostate cancer associated gene and biallelic markers  
 PT derived from it  
 XX  
 PS Claim 7; Page 251; 385pp; English.  
 XX  
 CC The invention relates to a mammalian PGI gene and protein, and a set of  
 CC PGI biallelic markers. The PGI polynucleotide and biallelic markers are  
 CC used in a hybridisation assay, a sequencing assay, or in an  
 CC allele-specific amplification assay for determining the identity of a  
 CC nucleotide at a PGI-related biallelic marker. The methods can be used to  
 CC detect and to assess the risk of developing cancer or prostate cancer.  
 CC Early-stage diagnosis of prostate cancer relies on prostate specific  
 CC antigen (PSA) dosage. However, the effectiveness of this is limited due  
 CC to its inability to discriminate between malignant and non-malignant  
 CC affections of the organ. A need exists for both a reliable diagnostic  
 CC procedure which would enable early-stage diagnosis, and for preventative  
 CC and curative treatments of the disease. The PGI gene can be used for  
 CC detection of prostate cancer, and the risk of developing it in the  
 CC future, and can also be used to determine therapies for the disease.  
 CC  
 XX  
 SQ Sequence 261 AA;

Query Match 56.4%; Score 199; DB 20; Length 261;  
 Best Local Similarity 100.0%; Pred. No. 1.9e-194;  
 Matches 199; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 155 MYLVIPEEGTRYNPEQTKVLSASQAFPAQRLAVLKHVLPRIKATHVAFDCMKNYDAI 214  
 Db 63 mylvifpeegtrrynpegtkvlisaqafaqr glavlkhvlprikatvafdcmknyda i 122  
 OY 215 YDVTVVVEGKDDGGRRESPTMEFLCECPKIHIDRIDKVPPEOEHRRLHREF 274  
 Db 123 ydvtvvvegkddggrresptmeflckeopk ihidridkvppeoge hmrtrwlherf 182  
 OY 275 EIKDKMLIEFESDPERRKRPFGKSVNSKLSIKRTLPMTLSGLTAGMIMTDAGRKLX 334  
 Db 183 eikdkmliefesdperrkrfp gksvnskslksrtlpmtllsgltagmimtdagr kly 242  
 OY 335 VNTWITGTLGCLWVTIKA 353  
 Db 243 vntwlygtllgclwvtika 261

RESULT 14  
 AAY36753  
 ID AAY36753 standard; Protein: 185 AA.  
 XX  
 AC AAY36753;  
 XX  
 DT 27-SEP-1999 (first entry)  
 XX  
 DE PGI splice variant allele protein sequence.  
 XX  
 KW PGI gene: biallelic marker; human; PSA; PGI-related biallelic marker;  
 KW cancer; prostate cancer; diagnosis; therapy; prostate specific antigen.  
 OS Homo sapiens.  
 XX  
 PN WO9932644-A2.  
 XX  
 PD 01-JUL-1999.  
 XX  
 PF 22-DEC-1998; 98WO-IB02133.

XX 09-SEP-1998; 98US-0099658.  
 PR 22-DEC-1997; 97US-0996306.  
 XX  
 XX (GEST ) GENSET.  
 PI Blumenfeld M, Bougueleret L, Chumakov I, Cohen D;  
 DR WPI; 1999-405178/34.  
 DR N-PSDB; AAZ00979.  
 XX  
 PT Use of a prostate cancer associated gene and biallelic markers  
 PT derived from it  
 PS Claim 7; Page 255; 385pp; English.  
 XX  
 CC The invention relates to a mammalian Pgl gene and protein, and a set of  
 CC Pgl biallelic markers. The Pgl polynucleotide and biallelic markers are  
 CC used in a hybridisation assay, a sequencing assay, or in an  
 CC allele-specific amplification assay for determining the identity of a  
 CC nucleotide at a Pgl-related biallelic marker. The methods can be used to  
 CC detect and to assess the risk of developing cancer or prostate cancer.  
 CC Early-stage diagnosis of prostate cancer relies on prostate specific  
 CC antigen (PSA) dosage. However, the effectiveness of this is limited due  
 CC to its inability to discriminate between malignant and non-malignant  
 CC affections of the organ. A need exists for both a reliable diagnostic  
 CC procedure which would enable early-stage diagnosis, and for preventative  
 CC and curative treatments of the disease. The Pgl gene can be used for  
 CC detection of prostate cancer, and the risk of developing it in the  
 CC future, and can also be used to determine therapies for the disease.  
 XX  
 SQ Sequence 185 AA;

Query Match 52.4%; Score 185; DB 20; Length 185;  
 Best Local Similarity 100.0%; Pred. No. 2.6e-180;  
 Matches 185; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 MRYLLPSVLLGTAPTVYVLAAGWWRLLSAFLPARFYQALDDRLCYQSWVLEFFENYTG 60  
 DB 1 MRYLLPSVLLGTAPTVYVLAAGWWRLLSAFLPARFYQALDDRLCYQSWVLEFFENYTG 60  
 QY 61 VQILLTGDELPKKNENITYLANHSTVDWIVADILAIROMALGHVRYLKGLKWLPLYGC 120  
 DB 61 VQILLTGDELPKKNENITYLANHSTVDWIVADILAIROMALGHVRYLKGLKWLPLYGC 120  
 QY 121 YFAOHGSIYVRSKAKFNEKEMRNKLSQSYVDAGTPMYLVIFPEGTRVNEPQTKVLSAQAF 180  
 DB 121 YFAOHGSIYVRSKAKFNEKEMRNKLSQSYVDAGTPMYLVIFPEGTRVNEPQTKVLSAQAF 180  
 QY 181 AAORG 185  
 DB 181 aaqrg 185

RESULT 15  
 ID AA36751  
 ID AA36751 standard; Protein; 315 AA.  
 AC AA36751;  
 XX  
 XX 27-SEP-1999 (first entry)  
 DT  
 XX  
 DE Pgl splice variant allele protein sequence.  
 XX  
 XX Pgl gene; biallelic marker; human; PSA; Pgl-related biallelic marker;  
 KW cancer; prostate cancer; diagnosis; therapy; prostate specific antigen.  
 XX  
 OS Homo sapiens.  
 XX  
 XX WO9932644-A2.  
 PN  
 XX 01-JUL-1999.  
 PD

XX 22-DEC-1998; 98WO-IB02133.  
 PF  
 XX  
 PR 09-SEP-1998; 98US-0099658.  
 PR 22-DEC-1997; 97US-0996306.  
 XX  
 XX (GEST ) GENSET.  
 PI Blumenfeld M, Bougueleret L, Chumakov I, Cohen D;  
 DR WPI; 1999-405178/34.  
 DR N-PSDB; AAZ00977.  
 XX  
 PT Use of a prostate cancer associated gene and biallelic markers  
 PT derived from it  
 PS Claim 7; Page 253-254; 385pp; English.  
 XX  
 CC The invention relates to a mammalian Pgl gene and protein, and a set of  
 CC Pgl biallelic markers. The Pgl polynucleotide and biallelic markers are  
 CC used in a hybridisation assay, a sequencing assay, or in an  
 CC allele-specific amplification assay for determining the identity of a  
 CC nucleotide at a Pgl-related biallelic marker. The methods can be used to  
 CC detect and to assess the risk of developing cancer or prostate cancer.  
 CC Early-stage diagnosis of prostate cancer relies on prostate specific  
 CC antigen (PSA) dosage. However, the effectiveness of this is limited due  
 CC to its inability to discriminate between malignant and non-malignant  
 CC affections of the organ. A need exists for both a reliable diagnostic  
 CC procedure which would enable early-stage diagnosis, and for preventative  
 CC and curative treatments of the disease. The Pgl gene can be used for  
 CC detection of prostate cancer, and the risk of developing it in the  
 CC future, and can also be used to determine therapies for the disease.  
 XX  
 SQ Sequence 315 AA;

Query Match 52.4%; Score 185; DB 20; Length 315;  
 Best Local Similarity 100.0%; Pred. No. 4.3e-180;  
 Matches 185; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 MRYLLPSVLLGTAPTVYVLAAGWWRLLSAFLPARFYQALDDRLCYQSWVLEFFENYTG 60  
 DB 1 MRYLLPSVLLGTAPTVYVLAAGWWRLLSAFLPARFYQALDDRLCYQSWVLEFFENYTG 60  
 QY 61 VQILLTGDELPKKNENITYLANHSTVDWIVADILAIROMALGHVRYLKGLKWLPLYGC 120  
 DB 61 VQILLTGDELPKKNENITYLANHSTVDWIVADILAIROMALGHVRYLKGLKWLPLYGC 120  
 QY 121 YFAOHGSIYVRSKAKFNEKEMRNKLSQSYVDAGTPMYLVIFPEGTRVNEPQTKVLSAQAF 180  
 DB 121 YFAOHGSIYVRSKAKFNEKEMRNKLSQSYVDAGTPMYLVIFPEGTRVNEPQTKVLSAQAF 180  
 QY 181 AAORG 185  
 DB 181 aaqrg 185

Search completed: August 28, 2002, 11:18:44  
 Job time: 582 sec

...



GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 28, 2002, 11:19:27 ; Search time 67.07 Seconds  
(without alignments)  
910.500 Million cell updates/sec

Title: US-09-853-526-4  
Perfect score: 353  
Sequence: 1 MRYLPSVLLGTAPTYLA.....YVNTWYGLTGLMTWTFKA 353

Scoring table:  
OLIGO  
Gapop 60.0 , Gapext 60.0

Searched: 562222 seqs, 172994929 residues

Word size : 0

Total number of hits satisfying chosen parameters: 562222

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database :

SPREMBL\_19:\*  
1: sp\_archaea:\*  
2: sp\_bacteria:\*  
3: sp\_fungi:\*  
4: sp\_human:\*  
5: sp\_invertebrate:\*  
6: sp\_mammal:\*  
7: sp\_mhc:\*  
8: sp\_organelle:\*  
9: sp\_phage:\*  
10: sp\_plant:\*  
11: sp\_rodent:\*  
12: sp\_virus:\*  
13: sp Vertebrate:\*  
14: sp Unclassified:\*  
15: sp\_rvirus:\*  
16: sp\_bacteriap:\*  
17: sp\_archaeap:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	353	100.0	364	4 Q9B0G4	Q9B0G4 homo sapien
2	252	71.4	353	4 Q9N0Q2	Q9N0Q2 homo sapien
3	32	9.1	354	11 Q9D1E8	Q9D1E8 mus musculu
4	9	2.5	264	16 Q9P0E7	Q9P0E7 ureaplasma
5	9	2.5	386	5 Q95R12	Q95R12 caenorhabdi
6	8	2.3	115	16 Q98E87	Q98E87 rhizobium l
7	8	2.3	236	16 Q9A675	Q9A675 caulobacter
8	8	2.3	262	16 Q84743	Q84743 chlamydia t
9	8	2.3	282	16 Q83380	Q83380 treponema p
10	8	2.3	411	16 Q9A4V3	Q9A4V3 caulobacter
11	8	2.3	430	2 Q54517	Q54517 actinobact
12	8	2.3	454	3 Q9P5W2	Q9P5W2 neuospora
13	8	2.3	458	16 Q9KFI0	Q9KFI0 bacillus su
14	8	2.3	459	16 Q31503	Q31503 bacillus su
15	8	2.3	483	10 Q49927	Q49927 pisum sativ
16	8	2.3	555	5 Q961R9	Q961R9 drosophila

## ALIGNMENTS

17	8	2.3	557	5 Q9VEX8	Q9VEX8 drosophila
18	8	2.3	1820	5 Q9VVG0	Q9VVG0 drosophila
19	7	2.0	54	6 Q28219	Q28219 cercopithec
20	7	2.0	78	16 Q98JCO	Q98JCO rhizobium l
21	7	2.0	126	11 Q62370	Q62370 mus musculu
22	7	2.0	132	11 Q9D581	Q9D581 mus musculu
23	7	2.0	137	12 Q55581	Q55581 leucania se
24	7	2.0	141	6 Q28218	Q28218 cercopithec
25	7	2.0	143	16 Q98RB9	Q98RB9 mycoplasma
26	7	2.0	148	2 Q9R25	Q9R25 streptomyce
27	7	2.0	148	16 Q9CH79	Q9CH79 lactococcus
28	7	2.0	150	2 Q9AGV0	Q9AGV0 bruceella ab
29	7	2.0	162	10 Q9LPP0	Q9LPP0 arabidopsis
30	7	2.0	163	2 Q9EX02	Q9EX02 streptomyce
31	7	2.0	175	2 Q9EX12	Q9EX12 klebsiella
32	7	2.0	184	11 Q9CWC3	Q9CWC3 mus musculu
33	7	2.0	195	16 Q9KCD7	Q9KCD7 bacillus ha
34	7	2.0	196	10 Q9LHL8	Q9LHL8 arabidopsis
35	7	2.0	209	16 Q910U7	Q910U7 pseudomonas
36	7	2.0	209	16 Q9CJKA	Q9CJKA pasteurella
37	7	2.0	211	16 Q67841	Q67841 aquifex ae
38	7	2.0	220	2 Q9EW7	Q9EW7 streptomyce
39	7	2.0	225	16 Q97NE9	Q97NE9 streptococ
40	7	2.0	230	16 Q9PH25	Q9PH25 campylobact
41	7	2.0	234	2 Q32330	Q32330 clostridium
42	7	2.0	240	12 Q9WHT2	Q9WHT2 thailand to
43	7	2.0	240	16 P73054	P73054 synchocyst
44	7	2.0	241	16 Q97KFA	Q97KFA clostridium
45	7	2.0	247	16 Q9XDL6	Q9XDL6 neisseria m

## RESULT 1

ID	Q9B0G4	PRELIMINARY;	PRT;	364 AA.
AC	Q9B0G4:			
DT	01-JUN-2001 (TREMBLrel. 17, Created)			
DT	01-DEC-2001 (TREMBLrel. 19, Last annotation update)			
DE	HYPOTHEICAL 42.1 KDA PROTEIN (LYSOPHOSPHATIDIC ACID			
DE	ACYLTRANSFERASE-EPSILON) (EC 2.3.1.51).			
GN	DKEP761C222 OR LPAAT-E.			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.			
OX	NCBI_Taxid=9606;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	TTSSUB-AMYGDAIA;			
RX	MDLINE-21154917; PubMed-11230166;			
RA	Wiemann S., Weil B., Wellenreuther R., Gaassenhuber J., Glassl S.,			
RA	Ansoerge W., Boecker M., Bloeker H., Bauersachs S., Blum H.,			
RA	Lauber J., Duesterhoeft A., Beyer A., Koehrer K., Strack N.,			
RA	Mewes H.W., Ottenwelder B., Obermaier B., Tampe J., Heubner D.,			
RA	Wambutt R., Korn B., Klein M., Poustka A.,			
RT	Towards a Catalog of Human Genes and Proteins: Sequencing and			
RT	Analysis of 500 Novel Complete Protein Coding Human cDNAs.";			
RL	Genome Res. 11:422-435(2001).			
RN	[2]			
RP	SEQUENCE FROM N.A.			
RA	Leung D.W.;			
RT	"Cloning and expression of LPAAT-epsilon.";			
RT	Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.			
DR	EMBL; AL136587; CAB66522.1; -			
DR	EMBL; AF375789; AAK54809.1; -			
DR	InterPro: IPR002123; Acyltransferase.			
DR	Pfam: PF01553; Acyltransferase; 1.			
KW	Hypothetical protein; Transferase; Acyltransferase.			
SQ	SEQUENCE 364 AA; 42072 MW; 90A0F87FC7C78081 CRC64;			

Query Match 100.0%; Score 353; DB 4; Length 364;  
 Best Local Similarity 100.0%; Pred. No. 0;  
 Matches 353; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MRYLLPSVLLGTAPTYVLAWGVRLLSAFLPARFYQALDRLCYVQSMVLEFFENTG 60  
 DB 12 MRYLLPSVLLGTAPTYVLAWGVRLLSAFLPARFYQALDRLCYVQSMVLEFFENTG 71  
 QY 61 VOILLYGDLPKKNENIYLANHGSTVDWIVADILAIRONALGHVRYVLEKGLKMLPLXGC 120  
 DB 72 VOILLYGDLPKKNENIYLANHGSTVDWIVADILAIRONALGHVRYVLEKGLKMLPLXGC 131  
 QY 121 YFAOHGCIYVRSKAFENKEKRNKQSYVDAGTPMYLVIPEEGTRYNEDEQTKVLSAQAF 180  
 DB 132 YFAOHGCIYVRSKAFENKEKRNKQSYVDAGTPMYLVIPEEGTRYNEDEQTKVLSAQAF 191  
 QY 181 AAOHGLAVLKHLVLPRIKATHVAFDCKMKNYIDATYDVTVYVEGDDGQRESPTMTPEFL 240  
 DB 192 AAOHGLAVLKHLVLPRIKATHVAFDCKMKNYIDATYDVTVYVEGDDGQRESPTMTPEFL 251  
 QY 241 CKECPKIHIDRIDKDVPEEOEHMRRLHEREFEIKDKMLIEFYESPDPERRRRFPCKS 300  
 DB 252 CKECPKIHIDRIDKDVPEEOEHMRRLHEREFEIKDKMLIEFYESPDPERRRRFPCKS 311  
 QY 301 VNSKLSIKTKLPMLISGLTAGMLMTDAGRKLVTWNTWYGTLLGCLMWTIKA 353  
 DB 312 VNSKLSIKTKLPMLISGLTAGMLMTDAGRKLVTWNTWYGTLLGCLMWTIKA 364

RESULT 2  
 ID Q9NUQ2 PRELIMINARY; PRT; 353 AA.  
 AC Q9NUQ2;  
 DT 01-OCT-2000 (Tremblrel. 15, Created)  
 DT 01-OCT-2000 (Tremblrel. 15, Last sequence update)  
 DT 01-JUN-2001 (Tremblrel. 17, Last annotation update)  
 DE CDNA FLJ11210 FTS, CLONE PLACE1007954.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.  
 NCBI\_Taxid=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=PLACENTA;  
 RA Isogai T., Ota T., Hayashi K., Sugiyama T., Otsuki T., Suzuki Y.,  
 RA Nishikawa T., Nagai K., Sugano S., Shiratori A., Sudo H.,  
 RA Magatsuma M., Hosoi T., Kaku Y., Kodaira H., Kondo H., Sugawara M.,  
 RA Takahashi M., Chiba Y., Ishida S., Murakawa K., Ono Y., Takiguchi S.,  
 RA Watanabe S., Kimura K., Murakami K., Ishii S., Kawai Y., Saito K.,  
 RA Yamamoto J., Wakamatsu A., Nakamura Y., Nagahari K., Masuno Y.,  
 RA Niimura Y., Iwayanagi T.;  
 RT "NEDO human cDNA sequencing project."  
 RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: AK002072; BAA92069.1; -;  
 DR InterPro: IPR002123; Acyltransferase.  
 DR Pfam: PF01553; Acyltransferase; 1.  
 SO SEQUENCE 353 AA; 40799 MW; 0ABAI4A2361F1B14 CRC64;

Query Match 71.4%; Score 252; DB 4; Length 353;  
 Best Local Similarity 99.7%; Pred. No. 3; Se-252;  
 Matches 352; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MRYLLPSVLLGTAPTYVLAWGVRLLSAFLPARFYQALDRLCYVQSMVLEFFENTG 60  
 DB 1 MRYLLPSVLLGTAPTYVLAWGVRLLSAFLPARFYQALDRLCYVQSMVLEFFENTG 60  
 QY 61 VOILLYGDLPKKNENIYLANHGSTVDWIVADILAIRONALGHVRYVLEKGLKMLPLXGC 120  
 DB 61 VOILLYGDLPKKNENIYLANHGSTVDWIVADILAIRONALGHVRYVLEKGLKMLPLXGC 120  
 QY 121 YFAOHGCIYVRSKAFENKEKRNKQSYVDAGTPMYLVIPEEGTRYNEDEQTKVLSAQAF 180  
 DB 121 YFAOHGCIYVRSKAFENKEKRNKQSYVDAGTPMYLVIPEEGTRYNEDEQTKVLSAQAF 180

DB 121 YFAOHGCIYVRSKAFENKEKRNKQSYVDAGTPMYLVIPEEGTRYNEDEQTKVLSAQAF 180  
 QY 181 AAOHGLAVLKHLVLPRIKATHVAFDCKMKNYIDATYDVTVYVEGDDGQRESPTMTPEFL 240  
 DB 181 AAOHGLAVLKHLVLPRIKATHVAFDCKMKNYIDATYDVTVYVEGDDGQRESPTMTPEFL 240  
 QY 241 CKECPKIHIDRIDKDVPEEOEHMRRLHEREFEIKDKMLIEFYESPDPERRRRFPCKS 300  
 DB 241 CKECPKIHIDRIDKDVPEEOEHMRRLHEREFEIKDKMLIEFYESPDPERRRRFPCKS 300  
 QY 301 VNSKLSIKTKLPMLISGLTAGMLMTDAGRKLVTWNTWYGTLLGCLMWTIKA 353  
 DB 301 VNSKLSIKTKLPMLISGLTAGMLMTDAGRKLVTWNTWYGTLLGCLMWTIKA 353

RESULT 3  
 ID Q9D1E8 PRELIMINARY; PRT; 354 AA.  
 AC Q9D1E8;  
 DT 01-JUN-2001 (Tremblrel. 17, Created)  
 DT 01-JUN-2001 (Tremblrel. 17, Last sequence update)  
 DT 01-JUN-2001 (Tremblrel. 17, Last annotation update)  
 DE 1110013A05RIK PROTEIN.  
 GN 1110013A05RIK.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 NCBI\_Taxid=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=C57BL/6J; TISSUE=EMBRYO;  
 RX MEDLINE=21085660; PubMed=11217851;  
 RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,  
 RA Arikawa T., Hara A., Fukunishi Y., Kono H., Adachi J., Fukuda S.,  
 RA Saito K., Ozaki Y., Gojobori T., Bono H., Katsukawa T., Saito R.,  
 RA Kaoto K., Matsuda H.A., Ashburner M., Batilov S., Casavant T.,  
 RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,  
 RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,  
 RA Schirml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,  
 RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,  
 RA Blake J., Boffelli D., Bojunga N., Carrinci P., de Bonaldo M.F.,  
 RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Kamiya M., Lee N.H.,  
 RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,  
 RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,  
 RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,  
 RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,  
 RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whitaker C., Wilming L.,  
 RA Wyshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohsaki S.,  
 RA Hayashizaki Y.;  
 RT "Functional annotation of a full-length mouse cDNA collection."  
 RL Nature 409:685-690(2001).  
 DR EMBL: AK003649; BAB22915.1; -;  
 DR MGD: MGI:1915880; 1110013A05RIK.  
 DR InterPro: IPR002123; Acyltransferase.  
 DR Pfam: PF01553; Acyltransferase; 1.  
 SO SEQUENCE 354 AA; 40943 MW; 447EEF924B91E800 CRC64;

Query Match 9.1%; Score 32; DB 11; Length 354;  
 Best Local Similarity 100.0%; Pred. No. 1; 9e-24;  
 Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 174 LSASQAFAPAGGLAVLKHLVLPRIKATHVARD 205  
 DB 174 LSASQAFAPAGGLAVLKHLVLPRIKATHVARD 205  
 RESULT 4  
 ID Q9POE7 PRELIMINARY; PRT; 264 AA.  
 AC Q9POE7;  
 DT 01-OCT-2000 (Tremblrel. 15, Created)

DT 01-OCT-2000 (TREMblrel. 15, last sequence update)  
DT 01-JUN-2001 (TREMblrel. 17, last annotation update)  
DE 1-ACYL-SN-GLYCEROL-3-PHOSPHATE ACYLTRANSFERASE.  
GN PLSC OR UU344.  
OS Ureaplasma parvum (Ureaplasma urealyticum biotype 1).  
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Mollicutes;  
OC Mycoplasmataceae; Ureaplasma.  
OX NCBI\_TaxID=134821;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-SEROVAR 3;  
RX MEDLINE=20500219; PubMed=11048724;  
RA Glass J.I., Lefkowitz E.J., Glass J.S., Heiner C.R., Chen E.Y.,  
RA Cassell G.H.;  
RT "The complete sequence of the mucosal pathogen Ureaplasma  
urealyticum";  
RL Nature 407:757-762(2000).  
DR EMBL; AF002131; AAF30753.1; -;  
DR InterPro; IPR002123; Acyltransferase.  
DR Pfam; PF01553; Acyltransferase; 1.  
KW Complete proteome.  
SQ SEQUENCE 264 AA; 30687 MW; DBC5727A07C7AF7 CRC64;

Query Match 2.5%; Score 9; DB 16; Length 264;  
Best Local Similarity 100.0%; Pred. No. 0.97;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 157 LVIFPEGTR 165  
DB 156 LVIFPEGTR 164

RESULT 5  
Q95R12 PRELIMINARY; PRT; 386 AA.  
AC Q95R12;  
DT 01-DEC-2001 (TREMblrel. 19, Created)  
DT 01-DEC-2001 (TREMblrel. 19, last sequence update)  
DT 01-DEC-2001 (TREMblrel. 19, last annotation update)  
DE HYPOTHETICAL 44.0 KDA PROTEIN.  
GN F28B3.9.  
OS Caenorhabditis elegans.  
OC Eukaryota; Metazoa; Nematozoa; Chromadorea; Rhabditida; Rhabditioidea;  
OC Rhabditidae; Peloderinae; Caenorhabditis.  
OX NCBI\_TaxID=6239;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-BRISTOL N2;  
RX MEDLINE=99069613; PubMed=9851916;  
RA None;  
RT "Genome sequence of the nematode C. elegans: a platform for  
investigating biology. The C. elegans Sequencing Consortium.";  
RL Science 282:2012-2018(1998).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN-BRISTOL N2;  
RA Geisel C., Kramer J., Smith A.;  
RT "The sequence of C. elegans cosmid F28B3.";  
RL Submitted (MAY-1997) to the EMBL/GenBank/DBJ databases.  
RN [3]  
RP SEQUENCE FROM N.A.  
RC STRAIN-BRISTOL N2;  
RA Waterston R.;  
RT "Direct Submission.";  
RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AF003136; AAK93853.1; -;  
KW Hypothetical protein.  
SQ SEQUENCE 386 AA; 43978 MW; AIE3020C201452A6 CRC64;

Query Match 2.5%; Score 9; DB 5; Length 386;  
Best Local Similarity 100.0%; Pred. No. 1.3;

Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 211 LDAIYDVT 219  
DB 230 LDAIYDVT 238

RESULT 6  
Q98E87 PRELIMINARY; PRT; 115 AA.  
AC Q98E87;  
DT 01-OCT-2001 (TREMblrel. 18, Created)  
DT 01-OCT-2001 (TREMblrel. 18, last sequence update)  
DT 01-OCT-2001 (TREMblrel. 18, last annotation update)  
DE MLL4358 PROTEIN.  
GN MLL4358.  
OS Rhizobium loti (Mesorhizobium loti).  
OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;  
OC Phyllobacteriaceae; Mesorhizobium.  
OX NCBI\_TaxID=381;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-MAFE30309;  
RX MEDLINE=21082930; PubMed=11214968;  
RA Kaneko T., Nakamura Y., Sato S., Asamizu E., Kato T., Sasamoto S.,  
RA Watanabe A., Idesawa K., Ishikawa A., Kawashima K., Kimura T.,  
RA Kishida Y., Kiyokawa C., Kohara M., Matsumoto M., Matsuno A.,  
RA Kouchizuki Y., Nakayama S., Nakazaki N., Shimpo S., Sugimoto M.,  
RA Takeuchi C., Yamada M., Tabata S.;  
RT "Complete genome structure of the nitrogen-fixing symbiotic bacterium  
Mesorhizobium loti.";  
RL DNA Res. 7:331-338(2000).  
DR EMBL; AP003004; BAB51033.1; -;  
KW Complete proteome.  
SQ SEQUENCE 115 AA; 12335 MW; 6AF43D6451F42BE8 CRC64;

Query Match 2.3%; Score 8; DB 16; Length 115;  
Best Local Similarity 100.0%; Pred. No. 5.1;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 174 LSASQAF 181  
DB 31 LSASQAF 38

RESULT 7  
Q9A675 PRELIMINARY; PRT; 236 AA.  
AC Q9A675;  
DT 01-JUN-2001 (TREMblrel. 17, Created)  
DT 01-JUN-2001 (TREMblrel. 17, last sequence update)  
DT 01-DEC-2001 (TREMblrel. 19, last annotation update)  
DE 1-ACYL-SN-GLYCEROL-3-PHOSPHATE ACYLTRANSFERASE.  
GN CC2219.  
OS Caulobacter crescentus.  
OC Bacteria; Proteobacteria; alpha subdivision; Caulobacter group;  
OC Caulobacter.  
OX NCBI\_TaxID=69394;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-ATCC 19089 / CB15;  
RX MEDLINE=21173698; PubMed=11259647;  
RA Nierman W.C., Feldblyum T.V., Laub M.T., Paulsen I.T., Nelson K.E.,  
RA Eisen J., Heidelberg J.F., Alley M.R.R., Ohta N., Maddock J.R.,  
RA Potocka I., Nelson W.C., Newton A., Stephens C., Phadke N.D., Ely B.,  
RA DeBoy R.T., Dosson R.J., Durkin A.S., Gwinn M.L., Haft D.H.,  
RA Kolonay J.F., Smit J., Craven M.B., Khouri H., Shetty J., Berry K.,  
RA Uitterback T., Tran K., Wolf A., Yamathavan J., Ernolaeva M., White O.,  
RA Salzberg S.L., Venter J.C., Shapiro L., Fraser C.M.;  
RT "Complete genome sequence of Caulobacter crescentus.";  
RL Proc. Natl. Acad. Sci. U.S.A. 98:4136-4141(2001).  
DR EMBL; AE005893; AAK24190.1; -;

DR TIGR; CC2219; -  
 DR InterPro; IPR002123; Acyltransferase.  
 DR Pfam; PF01553; Acyltransferase; 1.  
 KW Transferrase; Acyltransferase; Complete proteome.  
 SO SEQUENCE 236 AA; 26382 MW; 85694D3966BCA1A0 CRC64;

Query Match 2.3%; Score 8; DB 16; Length 236;  
 Best Local Similarity 100.0%; Pred. No. 9.5;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 OY 158 VIFPECTR 165  
 DB 146 VIFPECTR 153

RESULT 8  
 O84743 PRELIMINARY; PRT; 262 AA.  
 AC O84743;  
 DT 01-NOV-1998 (TREMblrel. 08, Created)  
 DT 01-NOV-1998 (TREMblrel. 08, Last sequence update)  
 DT 01-DEC-2001 (TREMblrel. 19, Last annotation update)  
 DE METAL-DEPENDENT HYDROLASE.  
 GN YJCU OR CT738.  
 OS Chlamydia trachomatis.  
 OC Bacteria; Chlamydiales; Chlamydiales; Chlamydia.  
 OX NCBI\_Taxid=813;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=D/WW-3/CX;  
 RX MEDLINE=99000809; PubMed=9784136;  
 RA Stephens R.S., Kalman S., Lammel C.J., Fan J., Marathe R., Aravind L.,  
 RA Mitchell W.P., Olinger L., Tatusov R.L., Zhao Q., Koonin E.V.,  
 RA Davis R.W.;  
 RT "Genome sequence of an obligate intracellular pathogen of humans:  
 RT Chlamydia trachomatis.";  
 RL Science 282:754-759(1998).  
 DR EMBL; AE001344; AAC68333.1; -;  
 KW Hydrolase; Complete proteome.  
 SO SEQUENCE 262 AA; 29082 MW; E36A6298A23CEFBF CRC64;

Query Match 2.3%; Score 8; DB 16; Length 262;  
 Best Local Similarity 100.0%; Pred. No. 10;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 OY 191 HVLPRPK 198  
 DB 202 HVLPRPK 209

RESULT 9  
 O83380 PRELIMINARY; PRT; 282 AA.  
 AC O83380;  
 DT 01-NOV-1998 (TREMblrel. 08, Created)  
 DT 01-NOV-1998 (TREMblrel. 08, Last sequence update)  
 DT 01-DEC-2001 (TREMblrel. 19, Last annotation update)  
 DE LYSOPHOSPHATIDIC ACID ACYLTRANSFERASE, PUTATIVE.  
 GN TP0361.  
 OS Treponema pallidum.  
 OC Bacteria; Spirochaetales; Spirochaetales; Treponema.  
 OX NCBI\_Taxid=160;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=NICHOLS;  
 RX MEDLINE=98332770; PubMed=965876;  
 RA Fraser C.M., Norris S.J., Weinstock G.M., White O., Sutton G.G.,  
 RA Dodson R., Gwin M., Hickey E.K., Clayton R., Ketchum K.A.,  
 RA Sodergren E., Hardham J.M., McLeod M.P., Salzberg S., Peterson J.,  
 RA Khalak H., Richardson D., Howell J.K., Chidambaram M., Utterback T.,  
 RA McDonald L., Artlich F., Bowman C., Colton M.D., Fujii C., Garland S.,

RA Hatch B., Horst K., Roberts K., Sandusky M., Weidman J., Smith H.O.,  
 RA Venter J.C.;  
 RT "Complete genome sequence of Treponema pallidum, the syphilis  
 RT spirochete.";  
 RL Science 281:375-388(1998).  
 DR EMBL; AE001215; AAC65346.1; -;  
 DR TIGR; TP0361;  
 DR InterPro; IPR002123; Acyltransferase.  
 DR Pfam; PF01553; Acyltransferase; 1.  
 KW Transferrase; Acyltransferase; Complete proteome.  
 SO SEQUENCE 282 AA; 31700 MW; DEF6903CD6C61AD CRC64;

Query Match 2.3%; Score 8; DB 16; Length 282;  
 Best Local Similarity 100.0%; Pred. No. 11;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 OY 158 VIFPECTR 165  
 DB 170 VIFPECTR 177

RESULT 10  
 O94AV3 PRELIMINARY; PRT; 411 AA.  
 AC O94AV3;  
 DT 01-JUN-2001 (TREMblrel. 17, Created)  
 DT 01-JUN-2001 (TREMblrel. 17, Last sequence update)  
 DT 01-DEC-2001 (TREMblrel. 19, Last annotation update)  
 DE METAL-ION EFFLUX MEMBRANE FUSION PROTEIN FAMILY.  
 GN CC2722.  
 OS Caulobacter crescentus.  
 OC Bacteria; Proteobacteria; alpha subdivision; Caulobacter group;  
 OC Caulobacter.  
 OX NCBI\_Taxid=69394;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=ATCC 19089 / CB15;  
 RX MEDLINE=21173698; PubMed=11259647;  
 RA Nierman W.C., Feldblyum T.V., Laub M.T., Paulsen I.T., Nelson K.E.,  
 RA Eisen J., Heidelberg J.F., Alley M.R.K., Ohta N., Maddock J.R.,  
 RA Potocka I., Nelson W.C., Newton A., Stephens C., Phadke N.D., Ely B.,  
 RA Deboy R.T., Dodson R.D., Durkin A.S., Gwin M.L., Haft D.H.,  
 RA Kolonay J.F., Smit J., Craven M.B., Khouri H., Shetty J., Berry K.,  
 RA Utterback T., Tran K., Wolf A., Yamathavan J., Ermolaeva M., White O.,  
 RA Salzberg S.L., Venter J.C., Shapiro L., Fraser C.M.;  
 RT "Complete genome sequence of Caulobacter crescentus.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 98:4136-4141(2001).  
 DR EMBL; AE005938; AAK24687.1; -;  
 DR TIGR; CC2722;  
 DR InterPro; IPR002215; HLYD.  
 DR Pfam; PF00529; HLYD; 1.  
 KW Complete proteome.  
 SO SEQUENCE 411 AA; 41818 MW; 1BA7849B9BCF22E CRC64;

Query Match 2.3%; Score 8; DB 16; Length 411;  
 Best Local Similarity 100.0%; Pred. No. 15;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 OY 316 ILSGLTAG 323  
 DB 380 ILSGLTAG 387

RESULT 11  
 O54517 PRELIMINARY; PRT; 430 AA.  
 AC O54517;  
 DT 01-JUN-1998 (TREMblrel. 06, Created)  
 DT 01-JUN-1998 (TREMblrel. 06, Last sequence update)  
 DT 01-OCT-2001 (TREMblrel. 18, Last annotation update)  
 DE 3-DEOXY-D-MANNO-2-OCTULOSONATE TRANSFERASE.

```
GN WAAA.
OS Acinetobacter baumannii.
OC Bacteria; Proteobacteria; gamma subdivision; Moraxellaceae;
OC Acinetobacter.
OX NCBL_TaxID=470;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-ATCC 19606, AND ATCC 15308;
RA Bode C.E., Brabetz W., Brade H.;
RL Submitted (JAN-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL; Z56926; CAB09652.1; -.
DR EMBL; Z56925; CAB09651.1; -.
KW Transferase.
SQ SEQUENCE 430 AA; 49070 MW; 695293748185D272 CRC64;

Query Match
Best Local Similarity 100.0%; Pred. No. 16;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 318 SGLTAGML 325
Db 171 SGLTAGML 178

RESULT 12
O9P5W2 PRELIMINARY; PRT; 454 AA.
ID O9P5W2;
AC O9P5W2;
DT 01-OCT-2000 (TREMBLrel. 15, Created)
DT 01-DEC-2001 (TREMBLrel. 19, last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, last annotation update)
DE RELATED TO GUANINE DEAMINASE.
GN B3E4.200.
OS Neurospora crassa.
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
OC Sordariales; Sordariaceae; Neurospora.
OX NCBL_TaxID=5141;
RN [1]
RP SEQUENCE FROM N.A.
RA Schulte U., Align V., Hoheisel J., Brandt P., Fartmann B., Holland R.,
RA Nyakatura G., Mewes H.W., Mannhaupt G.;
RL Submitted (MAY-2000) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA German Neurospora genome project;
RA Submitted (NOV-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AL355931; CAB91408.2; -.
SQ SEQUENCE 454 AA; 49778 MW; 8C1EC14BAECA421D CRC64;

Query Match
Best Local Similarity 100.0%; Pred. No. 17;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 189 LKHVTPR 196
Db 208 LKHVTPR 215

RESULT 13
O9KF10 PRELIMINARY; PRT; 458 AA.
ID O9KF10;
AC O9KF10;
DT 01-OCT-2000 (TREMBLrel. 15, Created)
DT 01-OCT-2000 (TREMBLrel. 15, last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, last annotation update)
DE RNA METHYLTRANSFERASE.
GN BH0687.
OS Bacillus halodurans.
OC Bacteria; Firmicutes; Bacillus/Clostridium group;
OC Bacillus/Staphylococcus group; Bacillus.
OX NCBL_TaxID=86665; RT
```

```
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-C-125 / JCM 9153;
RX MEDLINE=20512562; PubMed=11058132;
RA Takami H., Nakasone K., Takaki Y., Maeno G., Sasaki R., Masui N.,
RA Fuji F., Hirama C., Nakamura Y., Ogasawara N., Kuhara S.,
RA Horikoshi K.;
RT "Complete genome sequence of the alkaliphilic bacterium Bacillus
RT halodurans and genomic sequence comparison with Bacillus subtilis.";
RL Nucleic Acids Res. 28:4317-4331(2000).
DR EMBL; AP001509; BAB04406.1; -.
DR InterPro: IPR000051; SAM_bind.
DR InterPro: IPR001566; TRMA_1.
DR PROSITE: PS01230; TRMA_1; UNKNOWN_1.
KW Transferase; Methyltransferase; Complete proteome.
SQ SEQUENCE 458 AA; 51555 MW; 3D33A651E8B91240 CRC64;

Query Match
Best Local Similarity 100.0%; Pred. No. 17;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 167 NPEQTKVL 174
Db 291 NPEQTKVL 298

RESULT 14
O31503 PRELIMINARY; PRT; 459 AA.
ID O31503;
AC O31503;
DT 01-JAN-1998 (TREMBLrel. 05, Created)
DT 01-JAN-1998 (TREMBLrel. 05, last sequence update)
DT 01-OCT-2001 (TREMBLrel. 18, last annotation update)
DE YEFA PROTEIN.
GN YEFA.
OS Bacillus subtilis.
OC Bacteria; Firmicutes; Bacillus/Clostridium group;
OC Bacillus/Staphylococcus group; Bacillus.
OX NCBL_TaxID=1423;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-168;
RX MEDLINE=98044033; PubMed=9384377;
RA Kunst F., Ogasawara N., Moszer I., Albertini A.M., Alloni G.,
RA Azevedo V., Bertero M.G., Bessières P., Bolotin A., Borchert S.,
RA Borrijs R., Boursier L., Brans A., Braun M., Brignell S.C., Bron S.,
RA Brouillet S., Brusch C.V., Caldwell B., Capuano V., Carter N.M.,
RA Choi S.K., Codani J.J., Connerton I.F., Cummings N.J., Daniel R.A.,
RA Denizot F., Devine K.M., Dusterhoft A., Ehrlich S.D., Emerson P.T.,
RA Entlian K.D., Errington J., Fabret C., Ferrari E., Foulger D.,
RA Fritz C., Fujita M., Fujita Y., Funa S., Gallizzi A., Galleron N.,
RA Ghim S.Y., Glaser P., Goffeau A., Gollightly E.J., Grandi G.,
RA Giuseppe G., Guy B.J., Haga K., Haelech J., Harwood C.R., Henaut A.,
RA Hilbert H., Holsappel S., Hosono S., Hullo M.F., Itaya M., Jones L.,
RA Joris B., Karamata D., Kasahara Y., Klaerr-Bianchard M., Klein C.,
RA Kobayashi Y., Koelter P., Konigstein G., Krogh S., Kumano M.,
RA Kurita K., Lapidus A., Lardinois S., Lauber J., Lazarevic V.,
RA Lee S.M., Levine A., Liu H., Masuda S., Manuel C., Medigue C.,
RA Medina N., Mellado R.P., Mizuno M., Moestl D., Nakai S., Noback M.,
RA Noone D., O'Reilly M., Ogawa K., Ogiwara A., Oudega B., Park S.H.,
RA Parro V., Pohl T.M., Portetelle D., Porwollik S., Prescott A.M.,
RA Presecan E., Pujic P., Purnelle B., Rapoport G., Rey M., Reynolds S.,
RA Rieger M., Rivoita C., Rocha E., Roche B., Rose M., Sadale Y.,
RA Sato T., Scanlan E., Schleich S., Schroeter R., Scofield F.,
RA Sekiguchi J., Sekowska A., Seror S.J., Seror P., Shin B.S., Soldo B.,
RA Sorokin A., Taccani E., Takagi T., Takahashi H., Takemaru K.,
RA Takeuchi M., Tamakoshi A., Tanaka T., Terpstra P., Tognoni A.,
RA Tosato V., Uchiyama S., Vandenbol M., Vamlier F., Vassarotti A.,
RA Viari A., Wambutt R., Wedler E., Wedler H., Welzenegger T.,
RA Winters P., Wipat A., Yamamoto H., Yamane K., Yasunoto K., Yata K.,
RA Yoshida K., Yoshikawa H.F., Zumbstein E., Yoshikawa H., Danchin A.;
RT "The complete genome sequence of the gram-positive bacterium Bacillus
```

RT sub111s.":  
 RL Mature 390:249-256(1997).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-168:  
 RA Kunst F., Ogasawara N., Yoshikawa H., Danchin A.:  
 RL Submitted (NOV-1997) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; Z99107; CAB12493.1; -.  
 DR InterPro: IPR002792; DUF90.  
 DR InterPro: IPR000051; SAM\_bind.  
 DR InterPro: IPR001566; TRMA\_1.  
 DR Pfam: PF01938; TRAM; 1.  
 DR PROSITE: PS01230; TRMA\_1; UNKNOWN\_1.  
 DR PROSITE: PS01231; TRMA\_2; 1.  
 KM Complete proteome.  
 SQ SEQUENCE 459 AA; 51819 MW; 35E99547779421D5 CRC64;

Query Match 2.38; Score 8; DB 16; Length 459;  
 Best Local Similarity 100.0%; Pred. No. 17;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Caps 0;

OY 167 NPEQTKVL 174  
 |||||  
 DB 292 NPEQTKVL 299

RESULT 15  
 ID 049927 PRELIMINARY; PRT; 483 AA.  
 AC 049927;  
 DT 01-JUN-1998 (TREMBLrel. 06, Created)  
 DT 01-JUN-1998 (TREMBLrel. 06, Last sequence update)  
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)  
 DE P54. PROTEIN.  
 GN P54.  
 OS Pisum sativum (Garden pea).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
 OC eurosids I; Fabales; Fabaceae; Papilionoideae; Viciaeae; Pisum.  
 OX NCBI\_TaxID=3888;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Castillo J., Marquez J.A., Franco L., Ballestar E., Rodrigo M.I.;  
 RL Submitted (FEB-1997) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; Y11207; CAA72090.1; -.  
 DR HSSP: P50477; ICAU.  
 DR InterPro: IPR001113; Seedstore\_7s.  
 DR Pfam: PF00546; Seedstore\_7s; 1.  
 DR Pfam: PF02808; Seedstore\_7s\_C; 1.  
 SQ SEQUENCE 483 AA; 54662 MW; 8127BDAAA0178F3D CRC64;

Query Match 2.38; Score 8; DB 10; Length 483;  
 Best Local Similarity 100.0%; Pred. No. 18;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 151 AGTPMYLV 158  
 |||||  
 DB 184 AGTPMYLV 191

Search completed: August 28, 2002, 11:28:33  
 Job time: 546 sec

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 28, 2002, 10:59:36 ; Search time 75.55 Seconds

(without alignments)  
518.981 Million cell updates/sec

Title: US-09-853-526-4

Perfect score: 1871

Sequence: 1 MRVLLPSVILGTAFTYVLA.....YVNTWIVGTLGLMTWTKA 353

Scoring table:

BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 747574 seqs, 111073796 residues

Total number of hits satisfying chosen parameters: 747574

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database :

A.Geneseq\_032802.\*  
1: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1980.DAT.\*  
2: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1981.DAT.\*  
3: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1982.DAT.\*  
4: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1983.DAT.\*  
5: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1984.DAT.\*  
6: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1985.DAT.\*  
7: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1986.DAT.\*  
8: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1987.DAT.\*  
9: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1988.DAT.\*  
10: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1989.DAT.\*  
11: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1990.DAT.\*  
12: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1991.DAT.\*  
13: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1992.DAT.\*  
14: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1993.DAT.\*  
15: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1994.DAT.\*  
16: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1995.DAT.\*  
17: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1996.DAT.\*  
18: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1997.DAT.\*  
19: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1998.DAT.\*  
20: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1999.DAT.\*  
21: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA2000.DAT.\*  
22: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA2001.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	1871	100.0	353	22	AAB85532	Human secreted pro
2	1871	100.0	353	22	AAB85552	Human secreted pro
3	1871	100.0	364	20	AAV36729	Human PGI protein
4	1871	100.0	364	22	AAV50128	Human acyltransfer
5	1871	100.0	364	22	AAU00665	Human lysophosphat
6	1868	99.8	353	22	AAV33595	Human protein sequ
7	1867	99.8	352	20	AAV36728	Human PGI protein
8	1861	99.5	353	22	AAV39591	Human polypeptide
9	1860	99.4	372	22	AAV41377	Human polypeptide
10	1644	87.9	315	20	AAV36751	PGI splice variant
11	1568.5	83.8	354	20	AAV36741	Mouse PGI protein

12	1555.5	83.1	300	20	AAV36752	PGI splice variant
13	1498	80.1	291	20	AAV36744	PGI splice variant
14	1427	76.3	269	22	AAE09653	Human gene 15 enco
15	1325	70.8	261	20	AAV36745	PGI splice variant
16	1203	64.3	228	20	AAV36740	Truncated PGI prot
17	1182.5	63.2	238	20	AAV36743	PGI splice variant
18	976	52.2	185	20	AAV36753	PGI splice variant
19	820	43.8	182	20	AAV36750	PGI splice variant
20	449	24.0	96	20	AAV36749	PGI splice variant
21	401.5	21.5	378	21	AAV36749	Arabidopsis thalia
22	396.5	21.2	352	21	AAV36749	Arabidopsis thalia
23	393.5	21.0	341	21	AAV36749	Arabidopsis thalia
24	390.5	20.9	375	21	AAV36749	Arabidopsis thalia
25	390.5	20.9	375	21	AAV36749	Arabidopsis thalia
26	387.5	20.7	351	21	AAV36749	Arabidopsis thalia
27	387.5	20.7	375	21	AAV36749	Arabidopsis thalia
28	387	20.7	257	22	ABG01360	Novel human diagno
29	387	20.7	980	22	ABG11977	Novel human diagno
30	384.5	20.6	351	21	AAV36749	Arabidopsis thalia
31	374.5	20.0	343	21	AAV36749	Arabidopsis thalia
32	374.5	20.0	343	21	AAV36749	Arabidopsis thalia
33	332	17.7	77	20	AAV36749	PGI splice variant
34	330	17.6	66	20	AAV36749	PGI splice variant
35	329	17.6	68	20	AAV36749	PGI splice variant
36	329	17.6	90	20	AAV36749	PGI splice variant
37	328	17.5	374	22	AAU00667	Maize lysophosphat
38	326.5	17.5	374	21	AAU00667	Maize lysophosphat
39	326	17.4	374	15	AAV36749	Maize 2-acylttransf
40	315.5	16.9	374	21	AAV36749	Maize 2-acylttransf
41	311.5	16.6	387	21	AAV36749	Jojoba acyltransfe
42	309.5	16.6	377	17	AAV36749	Limnanthes CDNA cl
43	309.5	16.5	294	21	AAV36749	Zea mays protein I
44	308.5	16.5	389	21	AAV36749	Arabidopsis thalia
45	307.5	16.4	389	21	AAV36749	Arabidopsis thalia

## ALIGNMENTS

RESULT	1
ID	AAB85532 standard; protein; 353 AA.
AC	AAB85532;
XX	
DT	25-SEP-2001 (first entry)
XX	
DE	Human secreted protein (clone Id HUTH084).
XX	
KW	Secreted protein; immunosuppressive; antiarthritic; antirheumatic;
KW	antiproliferative; cytoskeletal; cardiac; vasotrophic; cerebroprotective;
KW	nootropic; neuroprotective; antibacterial; virucide; fungicide; human;
KW	ophthalmological; gene therapy.
OS	Homo sapiens.
XX	
PN	MO200155430-A1.
XX	
PD	02-AUG-2001.
XX	
PF	17-JAN-2001; 2001MO-US01431.
XX	
PR	31-JAN-2000; 2000US-0179065.
PR	04-FEB-2000; 2000US-0180628.
PR	12-SEP-2000; 2000US-0231968.
XX	
PA	(HUMA-) HUMAN GENOME SCI INC.
XX	
PI	Rosen CA, Komatsoulis GA, Baker KP, Birse CE, Soppet DR, Olsen HS;
PI	Moore PA, Wei P, Ehner R, Duan DR, Shi Y, Choi GH, Fiscella M;
PI	Ni J, Ruben SM, Barash SC;
XX	
DR	WPI; 2001-476220/51.

DR N-PSDB; AAH46942.

XX 17 Isolated nucleic acid molecules encoding human secreted proteins,  
PT used to preventing, treating or ameliorating a medical condition -

XX Claim 11: Page 450-451; 482pp; English.

XX The invention provides novel human secreted proteins and polynucleotides  
CC encoding them. The secreted proteins can be expressed by standard  
CC recombinant methodology. The secreted proteins and polynucleotides are  
CC used to prevent, treat or ameliorate a medical condition in e.g. humans,  
CC mice, rabbits, goats, horses, cats, dogs, chickens or sheep. They can  
CC also be used in diagnosing a pathological condition. The antibodies to  
CC the proteins can also be used in alleviating symptoms associated with the  
CC disorders and in diagnostic immunoassays e.g. radioimmunoassays or enzyme  
CC linked immunosorbent assays (ELISA). Disorders which are diagnosed or  
CC treated include autoimmune diseases e.g. rheumatoid arthritis,  
CC hyperproliferative disorders e.g. neoplasms of the breast or liver,  
CC cardiovascular disorders e.g. cardiac arrest, cerebrovascular disorders  
CC e.g. cerebral ischemia, angiodenesis, nervous system disorders e.g.  
CC Alzheimer's disease, infections caused by bacteria, viruses and fungi and  
CC ocular disorders e.g. corneal infection. The polypeptides can also be  
CC used to aid wound healing and epithelial cell proliferation, to prevent  
CC skin aging due to sunburn, to maintain organs before transplantation, for  
CC supporting cell culture of primary tissues, to regenerate tissues and in  
CC chemotaxis. The polypeptides can also be used as a food additive or  
CC preservative to increase or decrease storage capabilities. The present  
CC sequence represents a human secreted protein.

XX Sequence 353 AA:

Query Match 100.0%; Score 1871; DB 22; Length 353;

Best Local Similarity 100.0%; Pred. No. 2e-189;

Matches 353; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MRYLLPSVVLGTAFTYVLAAGVWRLLSAFLPARFYQALDDRLCYQSWLFFFEYNTG 60  
DB 1 mryllpsvvlgtapcyvlawgvrllsalfparfyqalddrlcyvqswmlfffeyntg 60  
QY 61 VOILLYGDLPRKKNENITYLANHSTVMDIVADLAIROMALGHVRVYLKEGKLPLRYGC 120  
DB 61 vqilllygdiprkkenillylanhgstvmdivadlalnqalnghvrlylkegklwplrygc 120  
QY 121 YFAOHGCIYVKSAAKNEKEMRNKLOSVDAGTPMYLVIFPEGSTRYNPEQTVLSASQAF 180  
DB 121 yfaohgciyvksaaknekeemrnlqsyvdagtpmylvifpegstryneqtkvlsasqaf 180  
QY 181 AAOGLAVLKHVLTPEIKATFVAFCMKNYLDAIDVTVVYEGKDDGOGRRSPMTTEFL 240  
DB 181 aagrlavlkhlvtpeikafvafcmknyldaidvltvvyeqgddgqrrspmttefl 240  
QY 241 CKECPKTHIHDIRDKRDVEDECHMRRLHERFEIKDKMLIEFYESPDERRRRPFGRS 300  
DB 241 ckecpkthihdirdkrdvdeechmrwlerfeikdkmliefyespderrrrpfgrs 300  
QY 301 VNSKLSIKTKLPSMLISGLTAGMLMTDAGRKLVTNTWIVGTLLGCLMTVITKA 353  
DB 301 vnsklstiktklpsmlisgltagmlmtdagrklyvntwivgtllgclmvtitka 353

# RESULT 2

AAH85552 2  
ID AAB85552 standard; protein; 353 AA.

AC AAB85552;

DT 25-SEP-2001 (first entry)

DE Human secreted protein (clone id HSLA181).

XX Secreted protein; immunosuppressive; antitubercic; antirheumatic;  
KM antiproliferative; cytostatic; cardiant; vasotropic; cerebroprotective;

KW nootropic; neuroprotective; antibacterial; virucide; fungicide; human;  
KM ophthalmological; gene therapy.

XX Homo sapiens.

XX WO20015430-A1.

XX 02-AUG-2001.

XX 17-JAN-2001; 2001WO-US01431.

XX 31-JAN-2000; 2000US-0179065.

XX 04-FEB-2000; 2000US-0180628.

XX 12-SEP-2000; 2000US-0231968.

XX (HUMA-) HUMAN GENOME SCI INC.

XX Rosen CA, Komatsoulis GA, Baker KP, Birse CE, Soppet DR, Olsen HS;  
PI Moore PA, Wei P, Ebner R, Duan DR, Shi Y, Choi GH, Fiscella M;  
PI Ni J, Ruben SM, Barash SC;

XX WPI: 2001-476220/51.

DR N-PSDB; AAH46962.

PT 17 Isolated nucleic acid molecules encoding human secreted proteins,  
used to preventing, treating or ameliorating a medical condition -

XX Claim 11: Page 465-466; 482pp; English.

XX The invention provides novel human secreted proteins and polynucleotides  
CC encoding them. The secreted proteins can be expressed by standard  
CC recombinant methodology. The secreted proteins and polynucleotides are  
CC used to prevent, treat or ameliorate a medical condition in e.g. humans,  
CC mice, rabbits, goats, horses, cats, dogs, chickens or sheep. They can  
CC also be used in diagnosing a pathological condition. The antibodies to  
CC the proteins can also be used in alleviating symptoms associated with the  
CC disorders and in diagnostic immunoassays e.g. radioimmunoassays or enzyme  
CC linked immunosorbent assays (ELISA). Disorders which are diagnosed or  
CC treated include autoimmune diseases e.g. rheumatoid arthritis,  
CC hyperproliferative disorders e.g. neoplasms of the breast or liver,  
CC cardiovascular disorders e.g. cardiac arrest, cerebrovascular disorders  
CC e.g. cerebral ischemia, angiodenesis, nervous system disorders e.g.  
CC Alzheimer's disease, infections caused by bacteria, viruses and fungi and  
CC ocular disorders e.g. corneal infection. The polypeptides can also be  
CC used to aid wound healing and epithelial cell proliferation, to prevent  
CC skin aging due to sunburn, to maintain organs before transplantation, for  
CC supporting cell culture of primary tissues, to regenerate tissues and in  
CC chemotaxis. The polypeptides can also be used as a food additive or  
CC preservative to increase or decrease storage capabilities. The present  
CC sequence represents a human secreted protein.

XX Sequence 353 AA:

Query Match 100.0%; Score 1871; DB 22; Length 353;

Best Local Similarity 100.0%; Pred. No. 2e-189;

Matches 353; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MRYLLPSVVLGTAFTYVLAAGVWRLLSAFLPARFYQALDDRLCYQSWLFFFEYNTG 60  
DB 1 mryllpsvvlgtapcyvlawgvrllsalfparfyqalddrlcyvqswmlfffeyntg 60  
QY 61 VOILLYGDLPRKKNENITYLANHSTVMDIVADLAIROMALGHVRVYLKEGKLPLRYGC 120  
DB 61 vqilllygdiprkkenillylanhgstvmdivadlalnqalnghvrlylkegklwplrygc 120  
QY 121 YFAOHGCIYVKSAAKNEKEMRNKLOSVDAGTPMYLVIFPEGSTRYNPEQTVLSASQAF 180  
DB 121 yfaohgciyvksaaknekeemrnlqsyvdagtpmylvifpegstryneqtkvlsasqaf 180  
QY 181 AAOGLAVLKHVLTPEIKATFVAFCMKNYLDAIDVTVVYEGKDDGOGRRSPMTTEFL 240  
DB 181 aagrlavlkhlvtpeikafvafcmknyldaidvltvvyeqgddgqrrspmttefl 240



QY 241 CKECPKIHIDRIDKKDVEDEQHMRRLHERFEIKDKMLIEYESPDERRKRFPFGKS 300  
|||||  
Db 241 ckecpkihndridkdkdveeqhmrtrw,herfeikdkmlieyespderrkrfpgks 300  
QY 301 VNSKLSIKTKLPMSMLILSGLTAGMLMTDAGRKLVTNTVIYGTLLGCLMTVITKA 353  
|||||  
Db 301 vnsklisiktklpmsmlilsgltagmlmtdagrklyvntwiygtllgclwvika 353  
RESULT 3  
AAY36729  
ID AAY36729 standard; Protein; 364 AA.  
XX  
AC AAY36729;  
XX  
DT 27-SEP-1999 (first entry)  
XX  
DE Human PGI protein sequence.  
XX  
KW PGI gene: diallelic marker; human; PSA: PGI-related diallelic marker;  
XX cancer; prostate cancer; diagnosis; therapy; prostate specific antigen.  
XX  
OS Homo sapiens.  
XX  
PN MO9932644-A2.  
XX  
PD 01-JUL-1999.  
XX  
PF 22-DEC-1998; 98WO-IB02133.  
XX  
PR 09-SEP-1998; 98US-0099658.  
XX 22-DEC-1997; 97US-0096306.  
XX  
PA (GEST ) GENSET.  
XX  
PI Blumenfeld M, Bougueleret L, Chumakov I, Cohen D;  
XX  
DR WPI: 1999-405178/34.  
XX  
DR N-PSDB; AA200870.  
XX  
PT Use of a prostate cancer associated gene and diallelic markers  
XX derived from it  
XX  
PS Claim 7; Page 190-191; 385pp; English.  
XX  
CC The invention relates to a mammalian PGI gene and protein, and a set of  
XX PGI diallelic markers. The PGI polynucleotide and diallelic markers are  
CC used in a hybridisation assay, a sequencing assay, or in an  
CC allele-specific amplification assay for determining the identity of a  
CC nucleotide at a PGI-related diallelic marker. The methods can be used to  
CC detect and to assess the risk of developing cancer or prostate cancer.  
CC Early-stage diagnosis of prostate cancer relies on prostate specific  
CC antigen (PSA) dosage. However, the effectiveness of this is limited due  
CC to its inability to discriminate between malignant and non-malignant  
CC affections of the organ. A need exists for both a reliable diagnostic  
CC procedure which would enable early-stage diagnosis, and for preventative  
CC and curative treatments of the disease. The PGI gene can be used for  
CC detection of prostate cancer, and the risk of developing it in the  
CC future, and can also be used to determine therapies for the disease.  
XX  
XX  
XX Sequence 364 AA;  
QY  
Query Match 100.0%; Score 1871; DB 20; Length 364;  
Best Local Similarity 100.0%; Pred. No. 2,1e-189;  
Matches 353; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Db 12 MYLLSVVLLGAPRYVLAAGWVRLSARLPARFQALDDRLXYCYQSVAFFEFYWTG 60  
1 MRVLLSVVLLGAPRYVLAAGWVRLSARLPARFQALDDRLXYCYQSVAFFEFYWTG 60  
Db 12 MYLLSVVLLGAPRYVLAAGWVRLSARLPARFQALDDRLXYCYQSVAFFEFYWTG 71  
1 MRVLLSVVLLGAPRYVLAAGWVRLSARLPARFQALDDRLXYCYQSVAFFEFYWTG 71  
QY 61 VQLLYGLDLPKKNENIYLANHOSVDMIVADILAIRQNALGHVRYVLRKGLKWLPLYGC 120

Db 72 VQALLYGLDLPKKNENIYLANHOSVDMIVADILAIRQNALGHVRYVLRKGLKWLPLYGC 131  
QY 121 YFAOHGIIYKRSKAFNEKEKMRNKLQSYVDAQPTMYIVITPEGRTRYVPEOTKYLSASQAF 180  
121 YFAOHGIIYKRSKAFNEKEKMRNKLQSYVDAQPTMYIVITPEGRTRYVPEOTKYLSASQAF 180  
Db 132 YFAGHGGIYKRSKAFNEKEKMRNKLQSYVDAQPTMYIVITPEGRTRYVPEOTKYLSASQAF 191  
132 YFAGHGGIYKRSKAFNEKEKMRNKLQSYVDAQPTMYIVITPEGRTRYVPEOTKYLSASQAF 191  
QY 181 AAORGGLVLRHVLPFRKATGVAFDCKKNYLDAYDTVYVEGKDDGQRRSPPTMEFL 240  
181 AAORGGLVLRHVLPFRKATGVAFDCKKNYLDAYDTVYVEGKDDGQRRSPPTMEFL 240  
Db 192 AAGQGLAVLRHVLPFRKATGVAFDCKKNYLDAYDTVYVEGKDDGQRRSPPTMEFL 251  
192 AAGQGLAVLRHVLPFRKATGVAFDCKKNYLDAYDTVYVEGKDDGQRRSPPTMEFL 251  
QY 241 CKECPKIHIDRIDKKDVEDEQHMRRLHERFEIKDKMLIEYESPDERRKRFPFGKS 300  
|||||  
Db 252 ckecpkihndridkdkdveeqhmrtrw,herfeikdkmlieyespderrkrfpgks 311  
252 ckecpkihndridkdkdveeqhmrtrw,herfeikdkmlieyespderrkrfpgks 311  
QY 301 VNSKLSIKTKLPMSMLILSGLTAGMLMTDAGRKLVTNTVIYGTLLGCLMTVITKA 353  
|||||  
Db 312 vnsklisiktklpmsmlilsgltagmlmtdagrklyvntwiygtllgclwvika 364  
312 vnsklisiktklpmsmlilsgltagmlmtdagrklyvntwiygtllgclwvika 364  
RESULT 4  
AAM50128  
ID AAM50128 standard; Protein; 364 AA.  
XX  
AC AAM50128;  
XX  
DT 21-DEC-2001 (first entry)  
XX  
DE Human acyltransferase 27417.  
XX  
KW Acyltransferase 27417; human; cell proliferation; cell migration;  
XX cell differentiation; ovary cancer; brain cancer; colon cancer;  
KW lung cancer; tumour; metastasis; sarcoma; carcinoma;  
XX adenocarcinoma; antitumour; lipid metabolism; diagnosis; therapy.  
XX  
XX Homo sapiens.  
XX  
FH Key  
FH Peptide  
FT /label= Signal\_peptide  
FT /label= 1..49  
FT /label= 50..364  
FT /label= Mature\_protein  
FT /label= 50..330  
FT /label= N-terminal\_domain  
FT /label= 321..337  
FT /label= Transmembrane\_domain  
FT /label= 338..364  
FT /label= C-terminal\_domain  
FT /label= 71..363  
FT Domain  
FT /label= Acyltransferase\_domain  
FT /label= 68..71  
FT /note= "asn is N-glycosylated"  
FT /label= 241..244  
FT /note= "O-phosphorylated by CAMP- and  
FT CGMP-dependent protein kinase"  
FT /label= 11..13  
FT /note= "O-phosphorylated by protein kinase C"  
FT /label= 144..146  
FT /note= "O-phosphorylated by protein kinase C"  
FT /label= 205..207  
FT /note= "O-phosphorylated by protein kinase C"  
FT /label= 317..319  
FT /note= "O-phosphorylated by protein kinase C"  
FT /label= 361..363  
FT /note= "O-phosphorylated by protein kinase C"  
FT /label= 95..98  
FT /note= "O-phosphorylated by casein kinase II"  
FT /label= 158..161  
FT /note= "O-phosphorylated by casein kinase II"  
FT /label= 246..249  
FT /note= "O-phosphorylated by casein kinase II"  
FT /label= 82..89  
FT /note= "O-phosphorylated by tyrosine kinase"

FT Modified-site 219..226  
 FT /note= "O-phosphorylated by tyrosine kinase"  
 FT Modified-site 23..28  
 FT /note= "N-myristoylated"  
 FT Modified-site 130..135  
 FT /note= "N-myristoylated"  
 FT Modified-site 330..335  
 FT /note= "N-myristoylated"  
 FT Modified-site 352..357  
 FT /note= "N-myristoylated"  
 FT Modified-site 230..343  
 FT /note= "Amidation"  
 FT Peptide 105..126  
 FT /note= "predicted leucine zipper"  
 XX MO200173051-A2.  
 XX 04-OCT-2001.  
 XX 26-MAR-2001; 2001WO-US09633.  
 XX 24-MAR-2000; 2000US-192092P.  
 XX (MILL-) MILLENNIUM PHARM INC.  
 XX Meyers RA, Rudolph-Owen L, Macbeth KJ, Tsai F, Williamson M;  
 XX MPI; 2001-626267/72.  
 XX N-PSDB; AAH27094.  
 DR  
 XX  
 PT New polynucleotide encodes a polypeptide for the diagnosis and  
 PT treatment of cancer and to alter cellular lipid metabolism comprises a  
 PT polynucleotide encoding human acyltransferases designated 46743 and  
 PT 27417  
 XX  
 XX Claim 9; Fig 8A-B; 150pp; English.  
 CC The present sequence is that of a novel human protein, termed  
 CC 27417, which shows the structural characteristics of members of the  
 CC acyltransferase family, including an acyltransferase domain. The  
 CC invention provides novel acyltransferase 27417 nucleic acids and  
 CC polypeptides, as well as methods for detecting their presence, and  
 CC methods for screening for compounds that modulate their expression  
 CC or activity. Such compounds can be used to treat conditions  
 CC related to aberrant activity or expression of 27417 protein or  
 CC nucleic acid, such as conditions involving deficient cellular  
 CC proliferation, migration and/or differentiation. Also provided are  
 CC methods for inhibiting the proliferation or migration, or inducing  
 CC the killing, of a 27417-expressing cell, e.g. a hyperproliferative  
 CC and/or metastatic cell, by contact with a compound that modulates  
 CC the activity or expression of the 27417 protein or nucleic acid.  
 CC The 27417-expressing cell is found in the healthy or diseased  
 CC heart, blood vessels, kidney, skeletal muscle, brain or liver, or  
 CC especially in a solid tumour, a soft tissue tumour or a metastatic  
 CC lesion, a sarcoma, a carcinoma or an adenocarcinoma, and in  
 CC particular in a hyperproliferative and/or metastatic cell found in  
 CC ovarian, brain, colon or lung cancer. The compounds may also be  
 CC used to modulate lipid metabolism in a 27417-expressing cell.  
 CC Methods of disease diagnosis, e.g. by determining the  
 CC presence of a genetic alteration in a 27417 polypeptide, and for  
 CC evaluating the efficacy of a treatment of a disorder, are also  
 CC provided.  
 CC  
 XX  
 XX Sequence 364 AA:  
 SQ  
 Query Match 100.0%; Score 1871; DB 22; Length 364;  
 Best Local Similarity 100.0%; Pred. No. 2.1e-189;  
 Matches 353; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 61 VOILLGDPKKNKIYLANHSTVDMIVADILAIRONALGHVRYVLEKGLKMLPLRYGC 120  
 |||||||  
 DB 72 VQILLGDPKKNKIYLANHSTVDMIVADILAIRONALGHVRYVLEKGLKMLPLRYGC 131  
 |||||||  
 QY 121 YFAOHGGLYVRSKFNENKRNKLSQYVDAGTMYLVIFPEGRYRNPDEQTKVLSASQAF 180  
 |||||||  
 DB 132 YFAHNGGLYVRSKFNENKRNKLSQYVDAGTMYLVIFPEGRYRNPDEQTKVLSASQAF 191  
 |||||||  
 QY 181 AAGRLAVLKHVLPRIKATHVAFDCKKNYDAIVDTVYVEGDDGORESPMTTEFL 240  
 |||||||  
 DB 192 AAGRLAVLKHVLPRIKATHVAFDCKKNYDAIVDTVYVEGDDGORESPMTTEFL 251  
 |||||||  
 QY 241 KCECPKIHIDRIDKKDVPEDQEHMRRLHREPEIKDKMLIEFYESPDERRRRPFGKS 300  
 |||||||  
 DB 252 KCECPKIHIDRIDKKDVPEDQEHMRRLHREPEIKDKMLIEFYESPDERRRRPFGKS 311  
 |||||||  
 QY 301 VNSKLSIKKTLPSMLISGLTAGMLMDAGKLVNVIYVITLGLCMLWRTKA 353  
 |||||||  
 DB 312 VNSKLSIKKTLPSMLISGLTAGMLMDAGKLVNVIYVITLGLCMLWRTKA 364  
 |||||||  
 RESULT 5  
 ID AAU00665 standard; Protein: 364 AA.  
 XX  
 XX AAU00665;  
 AC  
 DT 07-SEP-2001 (first entry)  
 XX  
 XX Human lysophosphatidic acid acyltransferase epsilon isoform LPAAT-epsilon.  
 DE  
 XX Lysophosphatidic acid acyltransferase epsilon; LPAAT-epsilon; LPA; PA;  
 KW lysophosphatidic acid; phosphatidic acid; acylation; cellular activation;  
 KW phospholipid signalling; mitogenesis; inflammation; autoimmune disease;  
 KW oncology; cancer; obesity; gene therapy.  
 XX  
 XX Homo sapiens.  
 OS  
 XX  
 XX MO200134782-A1.  
 XX  
 XX 17-MAR-2001.  
 PD  
 XX  
 XX 02-NOV-2000; 2000WO-US30193.  
 PF  
 XX  
 XX 09-NOV-1999; 99US-0436919.  
 PR  
 XX  
 XX (CELL-) CELL THERAPEUTICS INC.  
 PA  
 XX  
 XX Leung DW;  
 PT  
 XX  
 XX MPI; 2001-335920/35.  
 DR  
 XX  
 XX N-PSDB; AAS00649.  
 DR  
 XX  
 PT Novel isolated human isoform of lysophosphatidic acid  
 PT acyltransferase-epsilon useful for diagnostic, therapeutic and  
 PT screening purposes  
 PT  
 XX  
 PS Claim 1; Fig 1; 48pp; English.  
 CC The sequence represents a human lysophosphatidic acid acyltransferase  
 CC (LPAAT) isoform, LPAAT-epsilon. LPAAT catalyses the acylation of  
 CC lysophosphatidic acid (LPA) to phosphatidic acid (PA). LPA and PA have  
 CC been identified as phospholipid signalling molecules that affect a wide  
 CC range of biological responses. PA is involved in cellular activation and  
 CC mitogenesis. Compounds that block PA generation and hence diminish lipid  
 CC biosynthesis and the signal involved in cell activation and onco-  
 CC therapeutic interest in the areas of inflammation and oncology (e.g.  
 CC autoimmune diseases and cancer) as well as obesity treatment.  
 CC LPAAT-epsilon and its corresponding DNA can be used in screening assays  
 CC to detect agents that stimulate or inhibit the activity of LPAAT and,  
 CC therefore, PA. The DNA is useful in tests to detect the presence or  
 CC expression of LPAAT-epsilon in relation to certain diseases and

CC conditions, and in disease prevention and treatment. The sequences of the  
CC invention are also useful for diagnosis of diseases and conditions in  
CC which the expression of LPAAT enzyme is abnormal.  
XX

Sequence 364 AA:

Query Match 100.0%; Score 1871; DB 22; Length 364;  
Best Local Similarity 100.0%; Pred. No. 2.1e-189;  
Matches 353; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MRYLLPSVVLGTAPTYVLAAGVWRLLSAFLPAREYQALDDRLCYQSWLFEFFENYTG 60  
DB 12 MYLLPSVVLGTAPTYVLAAGVWRLLSAFLPAREYQALDDRLCYQSWLFEFFENYTG 71  
QY 61 VOILLGDLPRKKNENTITLANHSTVDWTVADILAIRONALGHVRYVLEKGLKWLPLXGC 120  
DB 72 VQILLGDLPRKKNENTITLANHSTVDWTVADILAIRONALGHVRYVLEKGLKWLPLXGC 131  
QY 121 YFAOHGCIYVKRSKAKFNEKEMRNKLOSVDAGTPMYLVIFPECTRRNPEQTKVLSAQA 180  
DB 132 YFAOHGCIYVKRSKAKFNEKEMRNKLOSVDAGTPMYLVIFPECTRRNPEQTKVLSAQA 191  
QY 181 AAQGLAVLKHLVLPRIKATHVAFCMKNYLDATYDVTVVYEGKDDGGRRESPTWTEFL 240  
DB 192 AAGRGLAVLKHLVLPRIKATHVAFCMKNYLDATYDVTVVYEGKDDGGRRESPTWTEFL 251  
QY 241 CKECPRIHIDRIKDDVPEDEHMRRLHEFEIKDKMLIEFYSPDERRRKRPFGKS 300  
DB 252 CKECPRIHIDRIKDDVPEDEHMRRLHEFEIKDKMLIEFYSPDERRRKRPFGKS 311  
QY 301 VNSKISIKKTLPSMLISGLTAGMLMTDAGRKLVTWTVYGTLLGCLWYTIKA 353  
DB 312 VNSKISIKKTLPSMLISGLTAGMLMTDAGRKLVTWTVYGTLLGCLWYTIKA 364

RESULT 6

AAB93595  
ID AAB93595 standard; Protein; 353 AA.

AC AAB93595;

DT 26-JUN-2001 (first entry)

DE Human protein sequence SEQ ID NO:13028.

KW Human; primer; detection; diagnosis; antisense therapy; gene therapy.

OS Homo sapiens.

PN EP1074617-A2.

PD 07-FEB-2001.

PE 28-JUL-2000; 2000EP-0116126.

PR 29-JUL-1999; 99JP-0248036.

PR 27-AUG-1999; 99JP-0300253.

PR 11-JAN-2000; 2000JP-0118776.

PR 02-MAY-2000; 2000JP-0183767.

PR 09-JUN-2000; 2000JP-0241899.

PA (HELI-) HELIX RES INST.

PI Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;

PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;

DR WPI; 2001-318749/34.

PT Primer sets for synthesizing polynucleotides, particularly the 5602  
PT full-length cDNAs defined in the specification, and for the detection  
PT and/or diagnosis of the abnormality of the proteins encoded by the  
PT full-length cDNAs -

XX Claim 8; SEQ ID 13028; 2537pp + CD ROW; English.

PS The present invention describes primer sets for synthesizing 5602  
XX full-length cDNAs defined in the specification. Where a primer set  
CC comprises: (a) an oligo-dr primer and an oligonucleotide complementary  
CC to the complementary strand of a polynucleotide which comprises one of  
CC the 5602 nucleotide sequences defined in the specification, where the  
CC oligonucleotide comprises at least 15 nucleotides; or (b) a combination  
CC of an oligonucleotide comprising a sequence complementary to the  
CC complementary strand of a polynucleotide which comprises a 5'-end  
CC sequence and an oligonucleotide comprising a sequence complementary to a  
CC polynucleotide which comprises a 3'-end sequence, where the  
CC oligonucleotide comprises at least 15 nucleotides and the combination of  
CC the 5'-end sequence/3'-end sequence is selected from those defined in  
CC the specification. The primer sets can be used in antisense therapy and  
CC in gene therapy. The primers are useful for synthesizing polynucleotides,  
CC particularly full-length cDNAs. The primers are also useful for the  
CC detection and/or diagnosis of the abnormality of the proteins encoded by  
CC the full-length cDNAs. The primers allow obtaining of the full-length  
CC cDNAs easily without any specialised methods. AAH03166 to AAH13628 and  
CC AAH13633 to AAH18742 represent human cDNA sequences; AAB92446 to  
CC AAB95893 represent human amino acid sequences; and AAH13629 to AAH13632  
CC represent oligonucleotides, all of which are used in the exemplification  
CC of the present invention.

Sequence 353 AA:

Query Match 99.8%; Score 1868; DB 22; Length 353;  
Best Local Similarity 99.7%; Pred. No. 4.2e-189;  
Matches 352; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 MRYLLPSVVLGTAPTYVLAAGVWRLLSAFLPAREYQALDDRLCYQSWLFEFFENYTG 60  
DB 1 MYLLPSVVLGTAPTYVLAAGVWRLLSAFLPAREYQALDDRLCYQSWLFEFFENYTG 60  
QY 61 VOILLGDLPRKKNENTITLANHSTVDWTVADILAIRONALGHVRYVLEKGLKWLPLXGC 120  
DB 61 VQILLGDLPRKKNENTITLANHSTVDWTVADILAIRONALGHVRYVLEKGLKWLPLXGC 120  
QY 121 YFAOHGCIYVKRSKAKFNEKEMRNKLOSVDAGTPMYLVIFPECTRRNPEQTKVLSAQA 180  
DB 121 YFAOHGCIYVKRSKAKFNEKEMRNKLOSVDAGTPMYLVIFPECTRRNPEQTKVLSAQA 180  
QY 181 AAQGLAVLKHLVLPRIKATHVAFCMKNYLDATYDVTVVYEGKDDGGRRESPTWTEFL 240  
DB 181 AAGRGLAVLKHLVLPRIKATHVAFCMKNYLDATYDVTVVYEGKDDGGRRESPTWTEFL 240  
QY 241 CKECPRIHIDRIKDDVPEDEHMRRLHEFEIKDKMLIEFYSPDERRRKRPFGKS 300  
DB 241 CKECPRIHIDRIKDDVPEDEHMRRLHEFEIKDKMLIEFYSPDERRRKRPFGKS 300  
QY 301 VNSKISIKKTLPSMLISGLTAGMLMTDAGRKLVTWTVYGTLLGCLWYTIKA 353  
DB 301 VNSKISIKKTLPSMLISGLTAGMLMTDAGRKLVTWTVYGTLLGCLWYTIKA 353

RESULT 7

AAV36728  
ID AAV36728 standard; Protein; 352 AA.

AC AAV36728;

DT 27-SEP-1999 (first entry)

DE Human Pgl protein sequence.

KW Pgl gene; diallelic marker; human; PSA; Pgl-related diallelic marker;  
KW cancer; prostate cancer; diagnosis; therapy; prostate specific antigen.

OS Homo sapiens.

PN WO932644-A2.  
 XX 01-JUL-1999.  
 PD  
 XX  
 PF 22-DEC-1998; 98MO-IB02133.  
 XX  
 PR 09-SEP-1998; 98US-0099658.  
 XX 22-DEC-1997; 97US-0996306.  
 XX  
 PA (GEST ) GENSET.  
 XX  
 PI Blumenfeld M, Bougueleret L, Chumakov I, Cohen D;  
 XX WPI: 1999-405178/34.  
 DR N-PSDB; AAZ00872.  
 XX  
 PT Use of a prostate cancer associated gene and biallelic markers  
 XX derived from it  
 PS  
 XX Claim 7; Page 190; 385pp; English.  
 XX  
 CC The invention relates to a mammalian Pgl gene and protein, and a set of  
 CC Pgl biallelic markers. The Pgl polynucleotide and biallelic markers are  
 CC used in a hybridisation assay, a sequencing assay, or in an  
 CC allele-specific amplification assay for determining the identity of a  
 CC nucleotide at a Pgl-related biallelic marker. The methods can be used to  
 CC detect and to assess the risk of developing cancer or prostate cancer.  
 CC Early-stage diagnosis of prostate cancer relies on prostate specific  
 CC antigen (PSa) dosage. However, the effectiveness of this is limited due  
 CC to its inability to discriminate between malignant and non-malignant  
 CC affections of the organ. A need exists for both a reliable diagnostic  
 CC procedure which would enable early-stage diagnosis, and for preventative  
 CC and curative treatments of the disease. The Pgl gene can be used for  
 CC detection of prostate cancer, and the risk of developing it in the  
 CC future, and can also be used to determine therapies for the disease.  
 CC  
 XX  
 SQ Sequence 352 AA:

Query Match 99.8%; Score 1867; DB 20; Length 352;  
 Best Local Similarity 100.0%; Pred. No. 5.4e-189;  
 Matches 352; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MRYLLPSVLLGTAPTYVLAMGWRLSAFLPARFYQALDRLCYQSWLFFENYTG 60  
 DB 1 mryllpsvllgtaptyvlamgwrlsafflparfyqalddrlcyqswmlffennytg 60  
 QY 61 VQILLYGDLPKKNKENTITLANHGSTVDWIVADILAIRONALGHVRYVLKGLKWLPLRYGC 120  
 DB 61 vqillygdldpkknkentiylanhgstvdwivadilairgnalghvryvlkqglkwlplygc 120  
 QY 121 YFAOHGIIYVRSKAFNEKEEMRNKLOSVDAGTPMYLVITPEGGRYRYPEQOTKVLASASQAF 180  
 DB 121 yfahogiiyvrsakfnekeemrnklqsyvdagtpmylvitfpeggrtynpeqtkvlasaqaf 180  
 QY 181 AAOGLAVLKLHVLTLPRIKATNVAFDCMKNYLDATYDVTVYVEGKDDGQGRRESPTMEFL 240  
 DB 181 aagrglavllkhlvltprlkathvafdcmknyldatydvtvlyegkddgqgrrespmtmefl 240  
 QY 241 CKECPRKHITIDRLDKDQVEBOEHMRRWYHERFEIKDKMLIEYESPDPERRKRRFGKS 300  
 DB 241 ckecpkhithidrldkdqvboehmrrwyherfelkdkmliefespdpertrkrrfpgks 300  
 QY 301 VNSKLSIKTKTLPMSLISGLTAGMLMTDAGRKLTVNTFWITGTLIGCLAWTIK 352  
 DB 301 vnsklstiktklpmslislsgltagmlmtdagrkltyvntfwitgtligclawtlk 352

RESULT 8  
 ID AAM39591 standard; protein; 353 AA.  
 XX  
 AC AAM39591;

XX 22-OCT-2001 (first entry)  
 DT  
 XX  
 DE Human polypeptide SEQ ID NO 2736.  
 XX  
 KW Human; nootropic; immunosuppressant; cytostatic; gene therapy; cancer;  
 KW peripheral nervous system; neuropathy; central nervous system; CNS;  
 KW Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;  
 KW amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;  
 KW chemokinetic; thrombolytic; drug screening; arthritis; inflammation;  
 KW leukaemia.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200153312-A1.  
 XX  
 PD 26-JUL-2001.  
 XX  
 PF 26-DEC-2000; 2000MO-US34263.  
 XX  
 PR 21-JAN-2000; 2000US-0488725.  
 PR 25-APR-2000; 2000US-0552317.  
 PR 09-JUL-2000; 2000US-0598042.  
 PR 19-JUL-2000; 2000US-0620312.  
 PR 03-AUG-2000; 2000US-0653450.  
 PR 14-SEP-2000; 2000US-0662191.  
 PR 19-OCT-2000; 2000US-0693036.  
 PR 29-NOV-2000; 2000US-0727344.  
 XX  
 PA (HYSE-) HYSEQ INC.  
 XX  
 XX Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;  
 PI Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J;  
 PI Zhao QA, Zhou P, Goodrich R, Drmanac RT;  
 XX  
 DR WPI: 2001-442253/47.  
 DR N-PSDB; AAI58747.  
 XX  
 PT Novel nucleic acids and polypeptides, useful for treating disorders  
 PT such as central nervous system injuries -  
 XX  
 PS Example 4; SEQ ID NO 2736; 10078bp; English.  
 XX  
 CC The invention relates to human nucleic acids (AA157798-AA161369) and  
 CC the encoded polypeptides (AAM38642-AAM42213) with nootropic,  
 CC immunosuppressant and cyostatic activity. The polynucleotides are useful  
 CC in gene therapy. A composition containing a polypeptide or polynucleotide  
 CC of the invention may be used to treat diseases of the peripheral nervous  
 CC system, such as peripheral nervous injuries, peripheral neuropathy and  
 CC localised neuropathies and central nervous system diseases, such as  
 CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic  
 CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the  
 CC utilisation of the activities such as: Immune system suppression.  
 CC Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic  
 CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,  
 CC assays for receptor activity, arthritis and inflammation, leukaemias and  
 CC C.N.S disorders.  
 CC Note: The sequence data for this patent did not form part of the printed  
 CC specification.  
 CC  
 XX  
 SQ Sequence 353 AA:

Query Match 99.5%; Score 1861; DB 22; Length 353;  
 Best Local Similarity 99.2%; Pred. No. 2.3e-188;  
 Matches 350; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 MRYLLPSVLLGTAPTYVLAMGWRLSAFLPARFYQALDRLCYQSWLFFENYTG 60  
 DB 1 mryllpsvllgtaptyvlamgwrlsafflparfyqalddrlcyqswmlffennytg 60  
 QY 61 VQILLYGDLPKKNKENTITLANHGSTVDWIVADILAIRONALGHVRYVLKGLKWLPLRYGC 120  
 DB 61 vqillygdldpkknkentiylanhgstvdwivadilairgnalghvryvlkqglkwlplygc 120

```

Db      61  vqlllygdipkknkniylanqstvdwivadlalnqnalghvryvlkeglkwlpllygc 120
Qy      121  YFAOHGIGYVKSASFENEKRNKLSQSYDAGTAPMVLVFFPGTRNPNPOTKVLSSQAF 180
Db      121  yfaohgiglyvksakfnekeemnklsqsydagtgmvlvffpgtrypnpektvlssqaf 180
Qy      181  AAORGLAVLKHVLTPEIKATHVAFDCMKNYLDIAIDVTVVEGKDGGORRSPPTTEFL 240
Db      181  aagrglavlkhvltprikathvafdcmknyldaidvtvtvyegkddggrresptntefl 240
Qy      241  CKECKRIHHIDRIKDDVPEDQEMHRRWLHERFEIKDKMLIEFYSPDERRRKRPGRKS 300
Db      241  ckecpkrihhidridkddvpeeqemhrrwlherfelkdkmliefespderrkrtrpgks 300
Qy      301  VNSKLSIKKTLPSMLILSGLTAGMLMTDAGRKLVTYVWTYGTLLGCLWTYTIKA 353
Db      301  vnsklisikktlpsmlllsgltagmlmtdagrklyvntwlygtllgclwtlrika 353

RESULT  9
AAM41377 standard; Protein; 372 AA.
AC      AAM41377;
AD      AAM41377;
DE      22-OCT-2001 (first entry)
XX      Human polypeptide SEQ ID NO 6308.
XX      Human; nocotropic; immunosuppressant; cytostatic; gene therapy; cancer;
XX      peripheral nervous system; neuropathy; central nervous system; CNS;
XX      Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
XX      amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;
XX      chemokine; thrombolytic; drug screening; arthritis; inflammation;
XX      leukaemia.
XX      Homo sapiens.
XX      PN      WO200153312-A1.
XX      PD      26-JUL-2001.
XX      PF      26-DEC-2000; 2000WO-US34263.
XX      PR      21-JAN-2000; 2000US-0488725.
XX      PR      25-APR-2000; 2000US-0552317.
XX      PR      09-JUL-2000; 2000US-0598042.
XX      PR      19-JUL-2000; 2000US-0620312.
XX      PR      03-AUG-2000; 2000US-0653450.
XX      PR      14-SEP-2000; 2000US-0662191.
XX      PR      19-OCT-2000; 2000US-0693036.
XX      PR      29-NOV-2000; 2000US-0727344.
XX      PA      (HXSE-) HXSEQ INC.
XX      PI      Tang YF, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;
XX      Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J;
XX      Zhao QH, Zhou P, Goodrich R, Drmanac RT;
XX      DR      WPI: 2001-442253/47.
XX      DR      N-PSDB; AA160533.
XX      PT      Novel nucleic acids and polypeptides, useful for treating disorders
XX      such as central nervous system injuries -
XX      Example 2; SEQ ID NO 6308; 10078bp; English.
XX      CC      The invention relates to human nucleic acids (AA157798-AA161369) and
XX      CC      the encoded polypeptides (AAM38642-AAM42213) with nocotropic,
XX      CC      immunosuppressant and cytostatic activity. The polynucleotides are useful
XX      CC      in gene therapy. A composition containing a polypeptide or polynucleotide
XX      CC      of the invention may be used to treat diseases of the peripheral nervous
XX      CC      system, such as peripheral nervous injuries, peripheral neuropathy and

```

```

CC      localised neuropathies and central nervous system diseases, such as
CC      Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
CC      lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
CC      utilisation of the activities such as: immune system suppression,
CC      Activin/inhibin activity, chemotactic/chemokine activity, haemostatic
CC      and thrombolytic activity, cancer diagnosis and therapy, drug screening,
CC      assays for receptor activity, arthritis and inflammation, leukaemias and
CC      C.N.S disorders.
CC      Note: The sequence data for this patent did not form part of the printed
CC      specification.
XX      SQ      Sequence      372 AA:
XX      Query Match      99.4%; Score 1860; DB 22; Length 372;
XX      Best Local Similarity 99.7%; Pred. No. 3, 2e-188;
XX      Matches 352; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Qy      1  MRYLPSVVLGTAPEYVLAAGVWRLLSAFLPAPYQALDDRLCYQSMVLEFFENYTG 60
Db      20  mryllpsvvlgtapeyvlaagvwrllsafflpafyqaldldrlcyqsmvleffenytg 79
Qy      61  VOILLYGDLPKKNKENTIIYANHOSTVDWIVADILAIKRONALGHVRYVLKEGLKWLPLYGC 120
Db      80  vqlllygdipkknkniylanqstvdwivadlalnqnalghvryvlkeglkwlpllygc 139
Qy      121  YFAOHGIGYVKSASFENEKRNKLSQSYDAGTAPMVLVFFPGTRNPNPOTKVLSSQAF 180
Db      140  yfaohgiglyvksakfnekeemnklsqsydagtgmvlvffpgtrypnpektvlssqaf 199
Qy      181  AAORGLAVLKHVLTPEIKATHVAFDCMKNYLDIAIDVTVVEGKDGGORRSPPTTEFL 240
Db      200  aagrglavlkhvltprikathvafdcmknyldaidvtvtvyegkddggrresptntefl 259
Qy      241  CKECKRIHHIDRIKDDVPEDQEMHRRWLHERFEIKDKMLIEFYSPDERRRKRPGRKS 300
Db      260  ckecpkrihhidridkddvpeeqemhrrwlherfelkdkmliefespderrkrtrpgks 319
Qy      301  VNSKLSIKKTLPSMLILSGLTAGMLMTDAGRKLVTYVWTYGTLLGCLWTYTIKA 353
Db      320  vnsklisikktlpsmlllsgltagmlmtdagrklyvntwlygtllgclwtlrika 372

RESULT  10
AAV36751 standard; Protein; 315 AA.
ID      AAV36751;
AC      AAV36751;
AD      AAV36751;
DE      27-SEP-1999 (first entry)
XX      PGI splice variant allele protein sequence.
XX      PGI gene; biallelic marker; human; PSA; PGI-related biallelic marker;
XX      cancer; prostate cancer; diagnosis; therapy; prostate specific antigen.
XX      Homo sapiens.
XX      PN      WO9932644-A2.
XX      PD      01-JUL-1999.
XX      PF      22-DEC-1998; 98WO-1B02133.
XX      PR      09-SEP-1998; 98US-0099658.
XX      PR      22-DEC-1997; 97US-0996306.
XX      PA      (GSEST ) GENSEST.
XX      PI      Blumenfeld M, Bougueleret L, Chumakov I, Cohen D;
XX      WPI: 1999-405178/34.
XX      DR      N-PSDB; AA200977.

```



KW PGI gene; biallelic marker; human; PSA; PGI-related biallelic marker;  
KW cancer; prostate cancer; diagnosis; therapy; prostate specific antigen.

XX Homo sapiens.

XX W09932644-A2.

XX 01-JUL-1999.

XX 22-DEC-1998; 98MO-IB02133.

XX 09-SEP-1998; 98US-0099658.

XX 22-DEC-1997; 97US-0996306.

XX (GEST ) GENSET.

XX Blumenfeld M, Bougueleret L, Chumakov I, Cohen D;

DR WPI; 1999-405178/34.

DR N-PSDB; AAZ00978.

XX Use of a prostate cancer associated gene and biallelic markers  
PT derived from it

PS Claim 7; Page 254-255; 385pp; English.

XX The invention relates to a mammalian PGI gene and protein, and a set of  
CC PGI biallelic markers. The PGI polynucleotide and biallelic markers are  
CC used in a hybridisation assay, a sequencing assay, or in an  
CC allele-specific amplification assay for determining the identity of a  
CC nucleotide at a PGI-related biallelic marker. The methods can be used to  
CC detect and to assess the risk of developing cancer or prostate cancer.  
CC Early-stage diagnosis of prostate cancer relies on prostate specific  
CC antigen (PSA) dosage. However, the effectiveness of this is limited due  
CC to its inability to discriminate between malignant and non-malignant  
CC affections of the organ. A need exists for both a reliable diagnostic  
CC procedure which would enable early-stage diagnosis, and for preventative  
CC and curative treatments of the disease. The PGI gene can be used for  
CC detection of prostate cancer, and the risk of developing it in the  
CC future, and can also be used to determine therapies for the disease.

XX Sequence 300 AA;

Query Match 83.1%; Score 1555.5; DB 20; Length 300;

Best Local Similarity 85.0%; Pred. No. 4.3e-156; Indels 53; Gaps 1;

Matches 300; Conservative 0; Mismatches 0; Indels 53; Gaps 1;

QY 1 MRYLLPSVVLGTAPTYVLAMGVWRLLSAFLPARYQALDRLCYOSMVLFFENYTG 60

DB 1 mryllpsvvlgtaptyvlamgvwrllsafiparfyalddrlycvgsmvllffenytg 60

QY 61 VOILLYGLDPKKNENIYLANHOSVVDWIVADILAIRONALGHVRYLKEGKMLPLYGC 120

DB 61 vqillyglldpkkenilylanhqvstvdwlvadilairqnalghvrylkegkmlplygc 120

QY 121 YFAOHGIVYKRSKAFNEKEMNKLOSVDAGTPMYLVTFPEGSTRYNPEOTVLSASQAF 180

DB 121 yfahgiglyvkrksafnekenmknklosyvdagtpmylvlfpegrtrypnpeqtkvlsasqaf 180

QY 181 AAORGLAVLKHVLPRIKATHVAFDCMKNYLDIVDVTVVYEGKDDGGRRSPMTTEFL 240

DB 181 aagr-----efl 187

QY 241 CKECPKIHIDRIDKKVPEEOEHMRRLHREFEIKDKMLIEFYESPDPERRKRFPKGS 300

DB 188 ckecpkhhidridkkvpeegemrrwlherfekdkmlliefespeperrkrffpgks 247

QY 301 VNSKSIKTKLPMSMLISGLTAGMLATDAGRKLYNTWITYGTLGCLWYTITA 353

DB 248 vnsklsiktklpmsmlisgltagmlatdagrklyntwitygtlglclwytika 300

RESULT 13

AAV36744 standard; Protein; 291 AA.

XX AAV36744;

XX 27-SEP-1999 (first entry)

XX PGI splice variant allele protein sequence.

XX PGI gene; biallelic marker; human; PSA; PGI-related biallelic marker;

XX cancer; prostate cancer; diagnosis; therapy; prostate specific antigen.

XX Homo sapiens.

XX W09932644-A2.

XX 01-JUL-1999.

XX 22-DEC-1998; 98MO-IB02133.

XX 09-SEP-1998; 98US-0099658.

XX 22-DEC-1997; 97US-0996306.

XX (GEST ) GENSET.

XX Blumenfeld M, Bougueleret L, Chumakov I, Cohen D;

DR WPI; 1999-405178/34.

DR N-PSDB; AAZ00970.

XX Use of a prostate cancer associated gene and biallelic markers  
PT derived from it

PS Claim 7; Page 250-251; 385pp; English.

XX The invention relates to a mammalian PGI gene and protein, and a set of  
CC PGI biallelic markers. The PGI polynucleotide and biallelic markers are  
CC used in a hybridisation assay, a sequencing assay, or in an  
CC allele-specific amplification assay for determining the identity of a  
CC nucleotide at a PGI-related biallelic marker. The methods can be used to  
CC detect and to assess the risk of developing cancer or prostate cancer.  
CC Early-stage diagnosis of prostate cancer relies on prostate specific  
CC antigen (PSA) dosage. However, the effectiveness of this is limited due  
CC to its inability to discriminate between malignant and non-malignant  
CC affections of the organ. A need exists for both a reliable diagnostic  
CC procedure which would enable early-stage diagnosis, and for preventative  
CC and curative treatments of the disease. The PGI gene can be used for  
CC detection of prostate cancer, and the risk of developing it in the  
CC future, and can also be used to determine therapies for the disease.

XX Sequence 291 AA;

Query Match 80.1%; Score 1498; DB 20; Length 291;

Best Local Similarity 82.4%; Pred. No. 5.1e-150; Indels 62; Gaps 1;

Matches 291; Conservative 0; Mismatches 0; Indels 62; Gaps 1;

QY 1 MRYLLPSVVLGTAPTYVLAMGVWRLLSAFLPARYQALDRLCYOSMVLFFENYTG 60

DB 1 mryllpsvvlgtaptyvlamgvwrllsafiparfyalddrlycvgsmvllffenytg 60

QY 61 VOILLYGLDPKKNENIYLANHOSVVDWIVADILAIRONALGHVRYLKEGKMLPLYGC 120

DB 61 v----- 61

QY 121 YFAOHGIVYKRSKAFNEKEMNKLOSVDAGTPMYLVTFPEGSTRYNPEOTVLSASQAF 180

DB 62 ---qnggilyvkrksafnekenmknklosyvdagtpmylvlfpegrtrypnpeqtkvlsasqaf 118

QY 181 AAORGLAVLKHVLPRIKATHVAFDCMKNYLDIVDVTVVYEGKDDGGRRSPMTTEFL 240

DB 119 aagrglavlkhvlprikathvafdcmknyldaydvttvvyegkddggrspmttefl 178





17-NOV-2000; 2000US-0249211.  
 PR 17-NOV-2000; 2000US-0249212.  
 PR 17-NOV-2000; 2000US-0249213.  
 PR 17-NOV-2000; 2000US-0249214.  
 PR 17-NOV-2000; 2000US-0249215.  
 PR 17-NOV-2000; 2000US-0249216.  
 PR 17-NOV-2000; 2000US-0249217.  
 PR 17-NOV-2000; 2000US-0249218.  
 PR 17-NOV-2000; 2000US-0249244.  
 PR 17-NOV-2000; 2000US-0249245.  
 PR 17-NOV-2000; 2000US-0249264.  
 PR 17-NOV-2000; 2000US-0249265.  
 PR 17-NOV-2000; 2000US-0249297.  
 PR 17-NOV-2000; 2000US-0249299.  
 PR 17-NOV-2000; 2000US-0249300.  
 PR 01-DEC-2000; 2000US-0250160.  
 PR 01-DEC-2000; 2000US-0250391.  
 PR 05-DEC-2000; 2000US-0251030.  
 PR 05-DEC-2000; 2000US-0251988.  
 PR 05-DEC-2000; 2000US-0256719.  
 PR 06-DEC-2000; 2000US-0251479.  
 PR 08-DEC-2000; 2000US-0251856.  
 PR 08-DEC-2000; 2000US-0251868.  
 PR 08-DEC-2000; 2000US-0251869.  
 PR 08-DEC-2000; 2000US-0251989.  
 PR 11-DEC-2000; 2000US-0254097.  
 PR 05-JAN-2001; 2001US-02559678.  
 XX  
 PA (HUMA-) HUMAN GENOME SCI INC.  
 PI Rosen CA, Barash SC, Ruben SM;  
 XX  
 DR WPI: 2001-457716/49.  
 DR N-PSDB: AAD16620.  
 XX  
 PT Isolated lipid metabolism polypeptide for screening to identify  
 PT antagonists and agonists that may enhance or block activities mediated  
 PT by lipid metabolism proteins and also for testing and detection e.g.  
 PT diagnosis -  
 XX  
 PS Claim 11; SEQ ID NO 47; 453pp; English.  
 XX  
 CC AAD16606-AAD16627 represent cDNAs corresponding to human lipid metabolism  
 CC protein genes, and AAE09639-AAE09660 represent the proteins they encode.  
 CC AAD16628-AAD16647 represent human lipid metabolism genomic DNAs. The  
 CC genes and their corresponding proteins are useful for preventing,  
 CC treating or ameliorating medical conditions such as immune disorders,  
 CC inflammatory disorders (e.g., conjunctivitis, glossitis), autoimmune  
 CC diseases (e.g., rheumatoid arthritis, systemic lupus erythematosus),  
 CC respiratory disorders (asthma, allergy), CNS disorders (e.g., Alzheimer's  
 CC disease, Parkinson's disease), haematopoietic disorders, cardiovascular  
 CC disorders (e.g., atherosclerosis, coronary disease), hyperlipidaemia,  
 CC hypolipidaemia, lipidosis, familial hypercholesterolaemia, Gaucher's  
 CC disease, Tay-Sachs disease, diabetes mellitus, cancer and other enzyme  
 CC abnormalities. They are used for the prevention of mental retardation  
 CC and the enlargement of spleen and liver. They are used in gene therapy  
 CC and antisense therapy. The present sequence represents a human lipid  
 CC metabolism protein of the invention.  
 CC Note: The sequence data for this patent did not form part of the printed  
 CC specification, but was obtained in electronic format directly from WIPO  
 CC at ftp.wipo.int/pub/publ/published\_pct\_sequences.  
 XX  
 XX Sequence 269 AA:

Query Match 76.3%; Score 1427; DB 22; Length 269;  
 Best Local Similarity 100.0%; Pred. No. 1.5e-142;  
 Matches 269; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 85 TWDWIVADILAIRONALGHVRYLVKLGKWLPLGYCYFAOHGIGYVRSKAKFNEKEMRNK 144  
 DB 1 tvdwivadilaironalghvrylvkleglkwlpdygcyfaqhgqgilykrsakfnekemrnk 60

QY 145 LQSYVDAGTPMYLVIFPEEGTRYNPEOTKYVLASQAFAAORGILAVLKHVLPRIKATNVAF 204  
 DB 61 lqsyvdagtpmylvifpeegtrynpeotkyvlasaqafaagqlavlkhwlprikatnvaf 120  
 QY 205 DCKKNYLDALDYVTYVEGKDDGGQRRSEPTMEFLCKECPKIHIDRIDKKDVEEOE 264  
 DB 121 dckknyldaldyvtvyegkddggqrrseptmeflckeckpkihidridkkdveepe 180  
 QY 265 HMRRWLHERREIKDKMLIERESPDPERRKRFCKSVNSKLSIKKTLPSMLISGLTAGM 324  
 DB 181 hmrwrlherreikdkmliefespdperrkrfpkavnsklsikktpsmllsgltagm 240  
 QY 325 LMTDAGRKLYVNTWYGTLLGCLMTWTKA 353  
 DB 241 lmtdagrklyvntwlygtllgclwvtika 269

RESULT 15  
 AAY36745  
 ID AAY36745 standard; Protein; 261 AA.  
 AC AAY36745;  
 XX  
 XX 27-SEP-1999 (first entry)  
 DT  
 XX  
 DE PGI splice variant allele protein sequence.  
 XX  
 XX PGI gene; biallelic marker; human; PSA: PGI-related biallelic marker;  
 KW cancer; prostate cancer; diagnosis; therapy; prostate specific antigen.  
 KM  
 XX Homo sapiens.  
 OS  
 XX  
 XX W0932644-A2.  
 PN  
 XX  
 PD 01-JUL-1999.  
 XX  
 PF 22-DEC-1998; 98WO-1B02133.  
 XX  
 PR 09-SEP-1998; 98US-0099658.  
 PR 22-DEC-1997; 97US-0996306.  
 XX  
 XX  
 PA (GEST ) GENSET.  
 XX  
 PI Blumenfeld M, Bougueleret L, Chumakov I, Cohen D;  
 XX  
 DR WPI: 1999-405178/34.  
 DR N-PSDB: AA200971.  
 XX

PT Use of a prostate cancer associated gene and biallelic markers  
 PT derived from it  
 XX  
 XX  
 XX Claim 7; Page 251; 385pp; English.  
 PS  
 XX

The invention relates to a mammalian PGI gene and protein, and a set of  
 CC PGI biallelic markers. The PGI polynucleotide and biallelic markers are  
 CC used in a hybridisation assay, a sequencing assay, or in an  
 CC allele-specific amplification assay for determining the identity of a  
 CC nucleotide at a PGI-related biallelic marker. The methods can be used to  
 CC detect and to assess the risk of developing cancer or prostate cancer.  
 CC Early-stage diagnosis of prostate cancer relies on prostate specific  
 CC antigen (PSA) dosage. However, the effectiveness of this is limited due  
 CC to its inability to discriminate between malignant and non-malignant  
 CC affections of the organ. A need exists for both a reliable diagnostic  
 CC procedure which would enable early-stage diagnosis, and for preventative  
 CC and curative treatments of the disease. The PGI gene can be used for  
 CC detection of prostate cancer, and the risk of developing it in the  
 CC future, and can also be used to determine therapies for the disease.

Query Match 70.8%; Score 1325; DB 20; Length 261;

Best Local Similarity 73.9%; Pred. No. 9,4e-132;  
Matches 261; Conservative 0; Mismatches 0; Indels 92; Gaps 1;

```
OY 1 MRYLLPSVVLGTAPTYVLAWGVRLSAFLPAREFYQALDDRLCYVQSMVLPFFENYTG 60
    |||||||
Db 1 mryllpsvvl|gtaptvylawgvwrl|saflparfyqalddr|lcyvqsmv|lffenytg 60
    ||
OY 61 VOILLYGDLPNNKENIYYLANHGSTVDMIVADILAIRONALGHVRYVLKEGLKMLPLYGC 120
    ||
Db 61 vq----- 62
OY 121 YFAGHGGIYVKRSKAFNEKEKRNKLOSVDAGTPMYLYIFPESTRYNPQTKVLSASQAF 180
    |||||||
Db 63 -----mylvlfpegrlrynpqtkvl|saqaf 88
OY 181 AAORGLAVLKHYLPRIKATHVAFDCMKKNYLDATYDVTVVEGKDDGQORESPPTMPEFL 240
    |||||||
Db 89 aagrglavlkhvl|prlkacvafdcmknyldaldvltvyegkddgqgrrtesp|umtefl 148
OY 241 CKECPKIHIDRIDKKDVPEOEHRMRWLHEREEIKDKMLIEFYESPDPERRRRFPQGS 300
    |||||||
Db 149 ckecpkih|hidrid|dkkdvp|eegehm|rrwl|herfel|dkdml|iefy|espdp|errrr|fpqgs 208
OY 301 VNSKLSIKKTLPSMLISGLTAGMLMTDAGRKLYVNTWYIYGTLLGCLMTVTKA 353
    |||||||
Db 209 vnsklsl|kclt|psm|l|sgltagm|mtc|dagrk|lyvntw|ygl|l|gcl|w|v|t|k|a 261
```

Search completed: August 28, 2002, 11:07:33  
Job time: 477 sec

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 28, 2002, 11:03:21 ; Search time 31 Seconds  
(without alignments)  
278.137 Million cell updates/sec

Title: US-09-853-526-4

Perfect score: 1871

Sequence: 1 MRYLPSVVLGTAPRYVLA.....YVNTWIXGTLGLWVTIKK 353

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 231628 seqs, 24425594 residues

Total number of hits satisfying chosen parameters: 231628

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database : Issued\_Patents\_AA.\*  
1: /cgn2\_6/ptodata/2/1aa/5A.COMB.pep:\*  
2: /cgn2\_6/ptodata/2/1aa/5B.COMB.pep:\*  
3: /cgn2\_6/ptodata/2/1aa/6A.COMB.pep:\*  
4: /cgn2\_6/ptodata/2/1aa/6B.COMB.pep:\*  
5: /cgn2\_6/ptodata/2/1aa/PCITUS.COMB.pep:\*  
6: /cgn2\_6/ptodata/2/1aa/backfill1es1.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1871	100.0	353	2	US-08-996-306-4
2	1871	100.0	353	4	US-09-338-907-4
3	1871	100.0	353	4	US-09-218-207-4
4	1871	100.0	364	2	US-08-996-306-5
5	1871	100.0	364	4	US-09-338-907-5
6	1871	100.0	364	4	US-09-218-207-5
7	1644	87.9	315	4	US-09-338-907-134
8	1644	87.9	315	4	US-09-218-207-134
9	1568.5	83.8	354	4	US-09-338-907-74
10	1568.5	83.8	354	4	US-09-218-207-74
11	1555.5	83.1	300	4	US-09-338-907-135
12	1555.5	83.1	300	4	US-09-218-207-135
13	1498	80.1	291	4	US-09-338-907-127
14	1498	80.1	291	4	US-09-218-207-127
15	1325	70.8	261	4	US-09-338-907-128
16	1325	70.8	261	4	US-09-218-207-128
17	1203	64.3	228	4	US-09-338-907-70
18	1203	64.3	228	4	US-09-218-207-70
19	1182.5	63.2	238	4	US-09-338-907-126
20	1182.5	63.2	238	4	US-09-218-207-126
21	976	52.2	185	4	US-09-338-907-136
22	976	52.2	185	4	US-09-218-207-136
23	820	43.8	182	4	US-09-338-907-133
24	820	43.8	182	4	US-09-218-207-133
25	449	24.0	97	4	US-09-338-907-132
26	449	24.0	97	4	US-09-218-207-132
27	332	17.7	77	4	US-09-338-907-125

28	332	17.7	77	4	US-09-218-207-125	Sequence 125, App
29	330	17.6	66	4	US-09-338-907-131	Sequence 131, App
30	330	17.6	66	4	US-09-218-207-131	Sequence 131, App
31	329	17.6	68	4	US-09-338-907-130	Sequence 130, App
32	329	17.6	68	4	US-09-218-207-130	Sequence 130, App
33	329	17.6	90	4	US-09-338-907-129	Sequence 129, App
34	329	17.6	90	4	US-09-218-207-129	Sequence 129, App
35	328	17.5	374	2	US-08-454-267-2	Sequence 2, Appl
36	328	17.5	374	2	US-08-941-319-2	Sequence 2, Appl
37	328	17.5	374	2	US-08-941-319-6	Sequence 6, Appl
38	328	17.5	374	2	US-08-941-319-6	Sequence 6, Appl
39	328	17.5	374	4	US-09-035-098-2	Sequence 2, Appl
40	328	17.5	374	4	US-09-035-098-6	Sequence 6, Appl
41	328	17.5	374	4	US-09-215-252-5	Sequence 5, Appl
42	268.5	14.4	378	4	US-09-215-252-17	Sequence 17, Appl
43	240	12.8	376	4	US-09-215-252-13	Sequence 13, Appl
44	239	12.8	295	2	US-08-454-267-7	Sequence 7, Appl
45	239	12.8	295	2	US-08-941-319-7	Sequence 7, Appl

#### ALIGNMENTS

RESULT 1  
US-08-996-306-4  
; Sequence 4, Application US/08996306  
; Patent No. 5945522  
; GENERAL INFORMATION:  
; APPLICANT: Cohen, Daniel  
; APPLICANT: Chumakov, Ilya  
; APPLICANT: Blumenfeld, Marta  
; APPLICANT: Bougueret, Lydie  
; TITLE OF INVENTION: Prostate cancer gene  
; NUMBER OF SEQUENCES: 68  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Knobbe, Martens, Olson & Bear  
; STREET: 501 West Broadway  
; CITY: San Diego  
; STATE: California  
; COUNTRY: USA  
; ZIP: 92101-3505  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy Disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: Win95  
; SOFTWARE: Word  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/996,306  
; FILING DATE:  
; CLASSIFICATION:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Israelson, Ned A.  
; REGISTRATION NUMBER: 29,655  
; REFERENCE/DOCKET NUMBER: GENSET.018A  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (619) 235-8550  
; TELEFAX: (619) 235-0176  
; INFORMATION FOR SEQ ID NO: 4:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 353 amino acids  
; TYPE: AMINO ACID  
; STRANDEDNESS: SINGLE  
; TOPOLOGY: LINEAR  
; MOLECULE TYPE: PROTEIN  
; ORIGINAL SOURCE:  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: potential Transmembrane helix  
; LOCATION: 1..33  
; IDENTIFICATION METHOD: Rao and Argos method  
; FEATURE:  
; NAME/KEY: potential Transmembrane helix  
; LOCATION: 4..20

IDENTIFICATION METHOD: Klein, Kanehisa and Delisi method  
FEATURE:  
NAME/KEY: potential Transmembrane helix  
LOCATION: 4..24  
IDENTIFICATION METHOD: Eisenberg,Schwarz,Komarony  
IDENTIFICATION METHOD: and Wall method  
FEATURE:  
NAME/KEY: potential N-myristoylation site  
LOCATION: 12  
IDENTIFICATION METHOD: prosite match  
FEATURE:  
NAME/KEY: potential Transmembrane helix  
LOCATION: 50..70  
IDENTIFICATION METHOD: Eisenberg,Schwarz,Komarony and Wall method  
FEATURE:  
NAME/KEY: potential N-glycosylation site  
LOCATION: 57  
IDENTIFICATION METHOD: prosite match  
FEATURE:  
NAME/KEY: potential Transmembrane helix  
LOCATION: 76..96  
IDENTIFICATION METHOD: Eisenberg,Schwarz,Komarony and Wall method  
FEATURE:  
NAME/KEY: potential Tyrosine kinase phosphorylation site  
LOCATION: 78  
IDENTIFICATION METHOD: prosite match  
FEATURE:  
NAME/KEY: potential Casein kinase II phosphorylation site  
LOCATION: 84  
IDENTIFICATION METHOD: prosite match  
FEATURE:  
NAME/KEY: potential Leucine zipper pattern  
LOCATION: 94..115  
IDENTIFICATION METHOD: prosite match  
FEATURE:  
NAME/KEY: potential N-myristoylation site  
LOCATION: 119  
IDENTIFICATION METHOD: prosite match  
FEATURE:  
NAME/KEY: potential Protein kinase C phosphorylation site  
LOCATION: 133  
IDENTIFICATION METHOD: prosite match  
FEATURE:  
NAME/KEY: potential Casein kinase II phosphorylation site  
LOCATION: 147  
IDENTIFICATION METHOD: prosite match  
FEATURE:  
NAME/KEY: potential Protein kinase C phosphorylation site  
LOCATION: 194  
IDENTIFICATION METHOD: prosite match  
FEATURE:  
NAME/KEY: potential Tyrosine kinase phosphorylation site  
LOCATION: 215  
IDENTIFICATION METHOD: prosite match  
FEATURE:  
NAME/KEY: potential Tyrosine sulfatation site  
LOCATION: 221  
IDENTIFICATION METHOD: prosite match  
FEATURE:  
NAME/KEY: potential CAMP- and cGMP-dependent protein kinase phosphorylation site  
LOCATION: 223  
IDENTIFICATION METHOD: prosite match  
FEATURE:  
NAME/KEY: potential Casein kinase II phosphorylation site  
LOCATION: 235  
IDENTIFICATION METHOD: prosite match  
FEATURE:  
NAME/KEY: potential Protein kinase C phosphorylation site  
LOCATION: 306  
IDENTIFICATION METHOD: prosite match  
FEATURE:  
NAME/KEY: potential Transmembrane helix  
LOCATION: 310..330

IDENTIFICATION METHOD: Eisenberg,Schwarz,Komarony and Wall method  
FEATURE:  
NAME/KEY: potential N-myristoylation site  
LOCATION: 319  
IDENTIFICATION METHOD: prosite match  
FEATURE:  
NAME/KEY: potential N-myristoylation site  
LOCATION: 323  
IDENTIFICATION METHOD: prosite match  
FEATURE:  
NAME/KEY: potential Amidation site  
LOCATION: 329  
IDENTIFICATION METHOD: prosite match  
FEATURE:  
NAME/KEY: potential Transmembrane helix  
LOCATION: 333..353  
IDENTIFICATION METHOD: Eisenberg,Schwarz,Komarony and Wall method  
FEATURE:  
NAME/KEY: potential N-myristoylation site  
LOCATION: 341  
IDENTIFICATION METHOD: prosite match  
FEATURE:  
NAME/KEY: potential Protein kinase C phosphorylation site  
LOCATION: 350  
IDENTIFICATION METHOD: prosite match  
US-08-996-306-4  
Query Match 100.0%; Score 1871; DB 2; Length 353;  
Best Local Similarity 100.0%; Pred. No. 1.9e-194;  
Matches 353; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 MRYLPSVVLGTAPTYVLAMGVRLISAFLPARYQALDDRLCYQOSNVLEFFENYTG 60  
DB 1 MRYLPSVVLGTAPTYVLAMGVRLISAFLPARYQALDDRLCYQOSNVLEFFENYTG 60  
QY 61 VOILLYGDLPKRNENIYYLANHOSYVDWIVADILAIRONALGHVRYVLEKGLKMLPYGC 120  
DB 61 VOILLYGDLPKRNENIYYLANHOSYVDWIVADILAIRONALGHVRYVLEKGLKMLPYGC 120  
QY 121 YFAOHGSIYVRSKAKFNEKEMRNKLOSVDAGPMTLVIFPESTRNPEOTKVLASQAF 180  
DB 121 YFAOHGSIYVRSKAKFNEKEMRNKLOSVDAGPMTLVIFPESTRNPEOTKVLASQAF 180  
QY 181 AAOGLAVLKHVLTLPRIKATHVAFDCMKNYIDAIVDVTVVYEGSKDGGORRESPTMTTEL 240  
DB 181 AAOGLAVLKHVLTLPRIKATHVAFDCMKNYIDAIVDVTVVYEGSKDGGORRESPTMTTEL 240  
QY 241 CKECPRKIHIDRIDKNDVPEOEHHMRMLHERFEIKDKMLIEFYSPDERKRRPGKS 300  
DB 241 CKECPRKIHIDRIDKNDVPEOEHHMRMLHERFEIKDKMLIEFYSPDERKRRPGKS 300  
QY 301 VNSKLSIKKTLPSMLISGLTAGMLMTDAGRKYVNTWYIGTLGLWTYTKA 353  
DB 301 VNSKLSIKKTLPSMLISGLTAGMLMTDAGRKYVNTWYIGTLGLWTYTKA 353  
RESULT 2  
US-09-338-907-4  
Sequence 4, Application US/09338907  
Patent No. 6265346  
GENERAL INFORMATION:  
APPLICANT: Cohen, Daniel  
APPLICANT: Blumenfeld, Marita  
APPLICANT: Ilya, Chumakov  
APPLICANT: Bougueleret, Lydie  
TITLE OF INVENTION: PROSTATE CANCER GENE  
FILE REFERENCE: GENSET.18CP1CP  
CURRENT APPLICATION NUMBER: US/09/338,907  
CURRENT FILING DATE: 1999-06-23  
EARLIER APPLICATION NUMBER: 08/996,306  
EARLIER FILING DATE: 1997-12-22  
EARLIER APPLICATION NUMBER: 60/099,658

EARLIER FILING DATE: 1998-09-09  
EARLIER APPLICATION NUMBER: 09/218,207  
EARLIER FILING DATE: 1998-12-22  
NUMBER OF SEQ ID NOS: 578  
SOFTWARE: Patent.pm  
SEQ ID NO 4  
LENGTH: 353  
TYPE: PRT  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: HELIX  
LOCATION: 1..33  
OTHER INFORMATION: Rao and Argos identification method, potential helix  
FEATURE:  
NAME/KEY: HELIX  
LOCATION: 4..20  
OTHER INFORMATION: Klein, Kanehisa and Delisi identification method, potential  
FEATURE:  
NAME/KEY: HELIX  
LOCATION: 4..24  
OTHER INFORMATION: Eisenberg, Schwarz, Komarony, Wall identification method, F  
FEATURE:  
NAME/KEY: MYRISTATE  
LOCATION: 12..16  
OTHER INFORMATION: Prosite match  
FEATURE:  
NAME/KEY: HELIX  
LOCATION: 50..70  
OTHER INFORMATION: Eisenberg, Schwarz, Komarony, Wall identification method, F  
FEATURE:  
NAME/KEY: CARBOHYD  
LOCATION: 57..59  
OTHER INFORMATION: Prosite match  
FEATURE:  
NAME/KEY: HELIX  
LOCATION: 76..96  
OTHER INFORMATION: Eisenberg, Schwarz, Komarony, Wall identification method, F  
FEATURE:  
NAME/KEY: PHOSPHORYLATION  
LOCATION: 78  
OTHER INFORMATION: potential Tyrosine kinase site, Prosite match  
FEATURE:  
NAME/KEY: PHOSPHORYLATION  
LOCATION: 84  
OTHER INFORMATION: potential caseine kinase II site, Prosite match  
FEATURE:  
NAME/KEY: SITE  
LOCATION: 94..115  
OTHER INFORMATION: potential leucine zipper site, Prosite match  
FEATURE:  
NAME/KEY: MYRISTATE  
LOCATION: 119..123  
OTHER INFORMATION: potential site, Prosite match  
FEATURE:  
NAME/KEY: PHOSPHORYLATION  
LOCATION: 133  
OTHER INFORMATION: potential protein kinase C, Prosite match  
FEATURE:  
NAME/KEY: PHOSPHORYLATION  
LOCATION: 147  
OTHER INFORMATION: potential caseine kinase II site, Prosite match  
FEATURE:  
NAME/KEY: PHOSPHORYLATION  
LOCATION: 194  
OTHER INFORMATION: potential protein kinase C, Prosite match  
FEATURE:  
NAME/KEY: PHOSPHORYLATION  
LOCATION: 215  
OTHER INFORMATION: potential Tyrosine kinase site, Prosite match  
FEATURE:  
NAME/KEY: SULFATATION  
LOCATION: 221  
OTHER INFORMATION: Prosite match

FEATURE:  
NAME/KEY: PHOSPHORYLATION  
LOCATION: 223  
OTHER INFORMATION: potential cAMP and cGMP dependant protein kinase site,  
FEATURE:  
NAME/KEY: PHOSPHORYLATION  
LOCATION: 235  
OTHER INFORMATION: potential caseine kinase II site, Prosite match  
FEATURE:  
NAME/KEY: PHOSPHORYLATION  
LOCATION: 306  
OTHER INFORMATION: potential protein kinase C, Prosite match  
FEATURE:  
NAME/KEY: HELIX  
LOCATION: 310..330  
OTHER INFORMATION: Eisenberg, Schwarz, Komarony, Wall identification method,  
FEATURE:  
NAME/KEY: MYRISTATE  
LOCATION: 319..323  
OTHER INFORMATION: Prosite match  
FEATURE:  
NAME/KEY: MYRISTATE  
LOCATION: 323..327  
OTHER INFORMATION: Prosite match  
FEATURE:  
NAME/KEY: AMIDATION  
LOCATION: 329  
OTHER INFORMATION: Prosite match  
FEATURE:  
NAME/KEY: HELIX  
LOCATION: 333..353  
OTHER INFORMATION: Eisenberg, Schwarz, Komarony, Wall identification method,  
FEATURE:  
NAME/KEY: MYRISTATE  
LOCATION: 341..345  
OTHER INFORMATION: Prosite match  
FEATURE:  
NAME/KEY: PHOSPHORYLATION  
LOCATION: 350  
OTHER INFORMATION: potential protein kinase C, Prosite match  
US-09-338-907-4  
Query Match 100.0%; Score 1871; DB 4; Length 353;  
Best Local Similarity 100.0%; Pred. No. 1.9e-194;  
Matches 353; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 MRYLLPSVVLGTAPTYVLAMGVWRLSAPLPARFYQALDDRLCYCYQSMVLEFFENYTG 60  
|||||  
DB 1 MRYLLPSVVLGTAPTYVLAMGVWRLSAPLPARFYQALDDRLCYCYQSMVLEFFENYTG 60  
QY 61 VOILLYGDLPKKNENIYLANHOSYVDWIVADILAIQNALGHVRYLKEGKMLPLYGC 120  
|||||  
DB 61 VOILLYGDLPKKNENIYLANHOSYVDWIVADILAIQNALGHVRYLKEGKMLPLYGC 120  
QY 121 YFAOHGGIYVKRSAPKNEKEMRNKLOSVDAGPMTLVLPBSIRNPEQTKVLSASQAF 180  
|||||  
DB 121 YFAOHGGIYVKRSAPKNEKEMRNKLOSVDAGPMTLVLPBSIRNPEQTKVLSASQAF 180  
QY 181 AAOGLAVLKHVLTPIRIKATHVAFDCMKNYLDAIYDVTVAVEGKDDGGORRESPTTEFL 240  
|||||  
DB 181 AAOGLAVLKHVLTPIRIKATHVAFDCMKNYLDAIYDVTVAVEGKDDGGORRESPTTEFL 240  
QY 241 CKECPKIHIIHIDRIDKDDVPEDEQEHMRRLHFEFEIKDKMLIEFYSPPERKRPGRKS 300  
|||||  
DB 241 CKECPKIHIIHIDRIDKDDVPEDEQEHMRRLHFEFEIKDKMLIEFYSPPERKRPGRKS 300  
QY 301 VNSKLSIKKTLPSMLTSLGTLGMLTDAGRKLYVTWITYGTLGLCTWTIRA 353  
|||||  
DB 301 VNSKLSIKKTLPSMLTSLGTLGMLTDAGRKLYVTWITYGTLGLCTWTIRA 353  
RESULT 3

```

US-09-218-207-4
: Sequence 4, Application us/09218207
: Patent No. 6346381
: GENERAL INFORMATION:
:   APPLICANT: Cohen, Daniel
:   APPLICANT: Blumenfeld, Marta
:   APPLICANT: Ilya, Chumakov
:   APPLICANT: Bougueleret, Lydie
:   TITLE OF INVENTION: Prostate cancer gene
:   FILE REFERENCE: GENSET.018CP1
:   CURRENT APPLICATION NUMBER: US/09/218.207
:   CURRENT FILING DATE: 1998-12-22
:   EARLIER APPLICATION NUMBER: 08/996,306
:   EARLIER FILING DATE: 1997-12-22
:   EARLIER APPLICATION NUMBER: 60/099,658
:   EARLIER FILING DATE: 1998-09-09
:   NUMBER OF SEQ ID NOS: 578
:   SOFTWARE: Patent.pm
:   SEQ ID NO 4
:   LENGTH: 353
:   TYPE: PRT
:   ORGANISM: Homo sapiens
:   FEATURE:
:     NAME/KEY: HELIX
:     LOCATION: 1..33
:   OTHER INFORMATION: Rao and Argos identification method, potential helix
:   FEATURE:
:     NAME/KEY: HELIX
:     LOCATION: 4..20
:   OTHER INFORMATION: Klein, Kanehisa and Delisi identification method, potential helix
:   FEATURE:
:     NAME/KEY: HELIX
:     LOCATION: 4..24
:   OTHER INFORMATION: Eisenberg, Schwartz, Komarony, Wall identification method, potential helix
:   FEATURE:
:     NAME/KEY: MYRISTATE
:     LOCATION: 12..16
:   OTHER INFORMATION: Prosite match
:   FEATURE:
:     NAME/KEY: HELIX
:     LOCATION: 50..70
:   OTHER INFORMATION: Eisenberg, Schwartz, Komarony, Wall identification method, potential helix
:   FEATURE:
:     NAME/KEY: CARBOHYD
:     LOCATION: 57..59
:   OTHER INFORMATION: Prosite match
:   FEATURE:
:     NAME/KEY: HELIX
:     LOCATION: 76..96
:   OTHER INFORMATION: Eisenberg, Schwartz, Komarony, Wall identification method, potential helix
:   FEATURE:
:     NAME/KEY: PHOSPHORYLATION
:     LOCATION: 78
:   OTHER INFORMATION: potential Tyrosine kinase site, Prosite match
:   NAME/KEY: PHOSPHORYLATION
:     LOCATION: 84
:   OTHER INFORMATION: potential caseine kinase II site, Prosite match
:   FEATURE:
:     NAME/KEY: SITE
:     LOCATION: 94..115
:   OTHER INFORMATION: potential leucine zipper site, Prosite match
:   FEATURE:
:     NAME/KEY: MYRISTATE
:     LOCATION: 119..123
:   OTHER INFORMATION: potential site, Prosite match
:   FEATURE:
:     NAME/KEY: PHOSPHORYLATION
:     LOCATION: 133
:   OTHER INFORMATION: potential protein kinase C, Prosite match
:   FEATURE:
:     NAME/KEY: PHOSPHORYLATION
:     LOCATION: 147

```

	OTHER INFORMATION: potential	caseine kinase II site,	Prosite match
FEATURE:			
NAME/KEY: PHOSPHORYLATION			
LOCATION: 194			
OTHER INFORMATION: potential	protein kinase C,	Prosite match	
FEATURE:			
NAME/KEY: PHOSPHORYLATION			
LOCATION: 215			
OTHER INFORMATION: potential	Tyrosine kinase site,	Prosite match	
FEATURE:			
NAME/KEY: SULFATATION			
LOCATION: 221			
OTHER INFORMATION: Prosite	match		
FEATURE:			
NAME/KEY: PHOSPHORYLATION			
LOCATION: 233			
OTHER INFORMATION: potential	CAMP and cGMP dependant protein kinase site,		
FEATURE:			
NAME/KEY: PHOSPHORYLATION			
LOCATION: 235			
OTHER INFORMATION: potential	caseine kinase II site,	Prosite match	
FEATURE:			
NAME/KEY: PHOSPHORYLATION			
LOCATION: 306			
OTHER INFORMATION: potential	protein kinase C,	Prosite match	
FEATURE:			
NAME/KEY: HELIX			
LOCATION: 310..330			
OTHER INFORMATION: Eisenberg, Schwarz, Komarony, Wall	identification method,		
FEATURE:			
NAME/KEY: MYRISTATE			
LOCATION: 319..323			
OTHER INFORMATION: Prosite	match		
FEATURE:			
NAME/KEY: MYRISTATE			
LOCATION: 323..327			
OTHER INFORMATION: Prosite	match		
FEATURE:			
NAME/KEY: AMIDATION			
LOCATION: 329			
OTHER INFORMATION: Prosite	match		
FEATURE:			
NAME/KEY: HELIX			
LOCATION: 333..353			
OTHER INFORMATION: Eisenberg, Schwarz, Komarony, Wall	identification method,		
FEATURE:			
NAME/KEY: MYRISTATE			
LOCATION: 341..345			
OTHER INFORMATION: Prosite	match		
FEATURE:			
NAME/KEY: PHOSPHORYLATION			
LOCATION: 350			
OTHER INFORMATION: potential	protein kinase C,	Prosite match	
US-09-218-207-4			
Query Match	100.0%; Score 1871; DB 4; Length 353;		
Best Local Similarity	100.0%; Pred. No.1,9e-194;		
Matches	353; Conservative 0; Mismatches 0; Indels 0; Gaps 0;		
QY	1 MRLLPSVVLGTAPTYVLANGWRLSAPLPAREFYALDRLCYVOSMYLFFENYTG	60	
Db	1 MRLLPSVVLGTAPTYVLANGWRLSAPLPAREFYALDRLCYVOSMYLFFENYTG	60	
QY	61 VOILLYGDLPRNKKEIITLANHGSTVDIMIVADILAIRONALGHRYYLKELKMLPLYGC	120	
Db	61 VOILLYGDLPRNKKEIITLANHGSTVDIMIVADILAIRONALGHRYYLKELKMLPLYGC	120	
QY	121 YEAOHGGLYYVRSAKFENEKEMRNKLQSYVDAGTTPMYLVIPPEGRYRNEOYTKVLSASQAF	180	
Db	121 YFAOHGGGLYYVRSAKFENEKEMRNKLQSYVDAGTTPMYLVIPPEGRYRNEOYTKVLSASQAF	180	
QY	181 AAORGLAALKHVLTAPRIKATHVAFAFDCKMKNYLDIAIVDTVTVEYEGDDGGQRESPTMTTEFL	240	

|||||  
Db 181 AAORGLAVLKHVLTPIRIKATHVAFDCMKNYLDAIDVTYVYEGKDDGGRRSSPTMTPEL 240  
Qy 241 CKECKRIHIHDIRIDKDVPEOEHRMRWLHERFEIKDKMLIEFYESPDPERRKRPPGKS 300  
|||||  
Db 241 CKECKRIHIHDIRIDKDVPEOEHRMRWLHERFEIKDKMLIEFYESPDPERRKRPPGKS 300  
Qy 301 VNSKLSIKKTLPSMLILSGLTAGMLMTDAGRKLVTYNTWITYGTLGCLMWTIKA 353  
|||||  
Db 301 VNSKLSIKKTLPSMLILSGLTAGMLMTDAGRKLVTYNTWITYGTLGCLMWTIKA 353

## RESULT 4

US-08-996-306-5  
; Sequence 5, Application US/08996306  
; Patent No. 5945522  
; GENERAL INFORMATION:  
; APPLICANT: Cohen, Daniel  
; APPLICANT: Chumakov, Ilya  
; APPLICANT: Blumenfeld, Marta  
; APPLICANT: Bougueleret, Lydie  
; TITLE OF INVENTION: Prostate cancer gene  
; NUMBER OF SEQUENCES: 68  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Knobbe, Martens, Olson & Bear  
; STREET: 501 West Broadway  
; CITY: San Diego  
; STATE: California  
; COUNTRY: USA  
; ZIP: 92101-3505  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy Disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: Win95  
; SOFTWARE: Word  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/996,306  
; FILING DATE:  
; CLASSIFICATION:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Israelsen, Ned A.  
; REGISTRATION NUMBER: 29,655  
; REFERENCE/DOCKET NUMBER: GENSET, 018A  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (619) 235-8550  
; TELEFAX: (619) 235-0176  
; INFORMATION FOR SEQ ID NO: 5:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 364 amino acids  
; TYPE: AMINO ACID  
; STRANDEDNESS: SINGLE  
; TOPOLOGY: LINEAR  
; MOLECULE TYPE: PROTEIN  
; ORIGINAL SOURCE:  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: potential protein  
; LOCATION: 1..364  
; US-08-996-306-5

Query Match 100.0%; Score 1871; DB 2; Length 364;  
Best Local Similarity 100.0%; Pred. No. 2e-194;  
Matches 353; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 MRYLLPSVVLGTAPTYVLANGVWRLLSAFLEPARFYQALDDRLXCYOSMVLFFPENYTG 60  
|||||  
Db 12 MRYLLPSVVLGTAPTYVLANGVWRLLSAFLEPARFYQALDDRLXCYOSMVLFFPENYTG 71  
Qy 61 VOILLYGDLPKKNENIITLANHOSYVDWIVADIATIRONALGHVRYVLEKGLKWLPLYGC 120  
|||||  
Db 72 VOILLYGDLPKKNENIITLANHOSYVDWIVADIATIRONALGHVRYVLEKGLKWLPLYGC 131

Qy 121 YFAOHGIIYVKSRAKFNEMKRNKLOSVDAGTPMYLVTFPEGTRYNPBOTVLSASQAF 180  
|||||  
Db 132 YFAOHGIIYVKSRAKFNEMKRNKLOSVDAGTPMYLVTFPEGTRYNPBOTVLSASQAF 191  
Qy 181 AAORGLAVLKHVLTPIRIKATHVAFDCMKNYLDAIDVTYVYEGKDDGGRRSSPTMTPEL 240  
|||||  
Db 192 AAORGLAVLKHVLTPIRIKATHVAFDCMKNYLDAIDVTYVYEGKDDGGRRSSPTMTPEL 251  
Qy 241 CKECKRIHIHDIRIDKDVPEOEHRMRWLHERFEIKDKMLIEFYESPDPERRKRPPGKS 300  
|||||  
Db 252 CKECKRIHIHDIRIDKDVPEOEHRMRWLHERFEIKDKMLIEFYESPDPERRKRPPGKS 311  
Qy 301 VNSKLSIKKTLPSMLILSGLTAGMLMTDAGRKLVTYNTWITYGTLGCLMWTIKA 353  
|||||  
Db 312 VNSKLSIKKTLPSMLILSGLTAGMLMTDAGRKLVTYNTWITYGTLGCLMWTIKA 364

## RESULT 5

US-09-338-907-5  
; Sequence 5, Application US/09338907  
; Patent No. 626546  
; GENERAL INFORMATION:  
; APPLICANT: Cohen, Daniel  
; APPLICANT: Blumenfeld, Marta  
; APPLICANT: Ilya, Chumakov  
; APPLICANT: Bougueleret, Lydie  
; TITLE OF INVENTION: PROSTATE CANCER GENE  
; FILE REFERENCE: GENSET, 18C1CP  
; CURRENT APPLICATION NUMBER: US/09/338,907  
; EARLIER FILING DATE: 1999-06-23  
; EARLIER APPLICATION NUMBER: 08/996,306  
; EARLIER FILING DATE: 1997-12-22  
; EARLIER APPLICATION NUMBER: 60/099,658  
; EARLIER FILING DATE: 1998-09-09  
; EARLIER APPLICATION NUMBER: 09/218,207  
; EARLIER FILING DATE: 1998-12-22  
; NUMBER OF SEQ ID NOS: 578  
; SOFTWARE: Patent.pm  
; SEQ ID NO 5  
; LENGTH: 364  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; US-09-338-907-5

Query Match 100.0%; Score 1871; DB 4; Length 364;  
Best Local Similarity 100.0%; Pred. No. 2e-194;  
Matches 353; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MRYLLPSVVLGTAPTYVLANGVWRLLSAFLEPARFYQALDDRLXCYOSMVLFFPENYTG 60  
|||||  
Db 12 MRYLLPSVVLGTAPTYVLANGVWRLLSAFLEPARFYQALDDRLXCYOSMVLFFPENYTG 71  
Qy 61 VOILLYGDLPKKNENIITLANHOSYVDWIVADIATIRONALGHVRYVLEKGLKWLPLYGC 120  
|||||  
Db 72 VOILLYGDLPKKNENIITLANHOSYVDWIVADIATIRONALGHVRYVLEKGLKWLPLYGC 131  
Qy 121 YFAOHGIIYVKSRAKFNEMKRNKLOSVDAGTPMYLVTFPEGTRYNPBOTVLSASQAF 180  
|||||  
Db 132 YFAOHGIIYVKSRAKFNEMKRNKLOSVDAGTPMYLVTFPEGTRYNPBOTVLSASQAF 191  
Qy 181 AAORGLAVLKHVLTPIRIKATHVAFDCMKNYLDAIDVTYVYEGKDDGGRRSSPTMTPEL 240  
|||||  
Db 192 AAORGLAVLKHVLTPIRIKATHVAFDCMKNYLDAIDVTYVYEGKDDGGRRSSPTMTPEL 251  
Qy 241 CKECKRIHIHDIRIDKDVPEOEHRMRWLHERFEIKDKMLIEFYESPDPERRKRPPGKS 300  
|||||  
Db 252 CKECKRIHIHDIRIDKDVPEOEHRMRWLHERFEIKDKMLIEFYESPDPERRKRPPGKS 311  
Qy 301 VNSKLSIKKTLPSMLILSGLTAGMLMTDAGRKLVTYNTWITYGTLGCLMWTIKA 353  
|||||  
Db 312 VNSKLSIKKTLPSMLILSGLTAGMLMTDAGRKLVTYNTWITYGTLGCLMWTIKA 364

```

RESULT 6
US-09-218-207-5
; Sequence 5, Application US/09218207
; Patent No. 6346381
; GENERAL INFORMATION:
; APPLICANT: Cohen, Daniel
; APPLICANT: Blumenfeld, Marta
; APPLICANT: Ilya, Chumakov
; APPLICANT: Bougueleret, Lydie
; TITLE OF INVENTION: Prostate cancer gene
; FILE REFERENCE: GENSET.018CPI
; CURRENT APPLICATION NUMBER: US/09/218, 207
; CURRENT FILING DATE: 1998-12-22
; EARLIER APPLICATION NUMBER: 08/996,306
; EARLIER FILING DATE: 1997-12-22
; EARLIER APPLICATION NUMBER: 60/099,658
; EARLIER FILING DATE: 1998-09-09
; NUMBER OF SEQ ID NOS: 578
; SOFTWARE: Patent.pm
; SEQ ID NO 5
; LENGTH: 364
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-218-207-5

```

```

Query Match 100.0%; Score 1871; DB 4; Length 364;
Best Local Similarity 100.0%; Pred. No. 2e-194;
Matches 333; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1 MRYLLPSVLLGTAPTYVVLAMGWRLLSAFLPARFYQALDDRLCYVQSWLFEFFENYTG 60
DB 12 MRYLLPSVLLGTAPTYVVLAMGWRLLSAFLPARFYQALDDRLCYVQSWLFEFFENYTG 71
QY 61 VOILLGDDLPKKNENIYYLANHSTYDWMYADILAIROMALGHVRYVLEKGLMPLXGC 120
DB 72 VOILLGDDLPKKNENIYYLANHSTYDWMYADILAIROMALGHVRYVLEKGLMPLXGC 131
QY 121 YFAOHGIIYKRSKAKNEKEMRNKLOSVDAGTPMYLVIFPESTRYNEPTKVLASQAF 180
DB 132 YFAOHGIIYKRSKAKNEKEMRNKLOSVDAGTPMYLVIFPESTRYNEPTKVLASQAF 191
QY 181 MAORGLAVLKHVLTPIKATHVAFDCMKNYLDAIYDVTYVEKDGQGRRESPTMEFL 240
DB 192 MAORGLAVLKHVLTPIKATHVAFDCMKNYLDAIYDVTYVEKDGQGRRESPTMEFL 251
QY 241 CKCEPKIHIDRIKDDVPEBOEHMRWLHERFEIKDKMLIEFYSPDERRRKRPFGKS 300
DB 252 CKCEPKIHIDRIKDDVPEBOEHMRWLHERFEIKDKMLIEFYSPDERRRKRPFGKS 311
QY 301 VNSKLSIKKTLPMSMLISGLTAGMLMTDAGRKLVTWYIGTLGCLMWTIKA 353
DB 312 VNSKLSIKKTLPMSMLISGLTAGMLMTDAGRKLVTWYIGTLGCLMWTIKA 364

```

```

RESULT 7
US-09-338-907-134
; Sequence 134, Application US/09338907
; Patent No. 6265546
; GENERAL INFORMATION:
; APPLICANT: Cohen, Daniel
; APPLICANT: Blumenfeld, Marta
; APPLICANT: Ilya, Chumakov
; APPLICANT: Bougueleret, Lydie
; TITLE OF INVENTION: PROSTATE CANCER GENE
; FILE REFERENCE: GENSET.18CPI
; CURRENT APPLICATION NUMBER: US/09/338, 907
; CURRENT FILING DATE: 1999-06-23
; EARLIER APPLICATION NUMBER: 08/996,306
; EARLIER FILING DATE: 1997-12-22
; EARLIER APPLICATION NUMBER: 60/099,658
; EARLIER FILING DATE: 1998-09-09

```

```

; EARLIER APPLICATION NUMBER: 09/218,207
; EARLIER FILING DATE: 1998-12-22
; NUMBER OF SEQ ID NOS: 578
; SOFTWARE: Patent.pm
; SEQ ID NO 134
; LENGTH: 315
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: 81..83
; OTHER INFORMATION: Box 1
; FEATURE:
; NAME/KEY: SITE
; LOCATION: 160..165
; OTHER INFORMATION: Box 11
US-09-338-907-134

```

```

Query Match 87.9%; Score 1644; DB 4; Length 315;
Best Local Similarity 89.2%; Pred. No. 6.5e-170;
Matches 315; Conservative 0; Mismatches 0; Indels 38; Gaps 1;

```

```

QY 1 MRYLLPSVLLGTAPTYVVLAMGWRLLSAFLPARFYQALDDRLCYVQSWLFEFFENYTG 60
DB 1 MRYLLPSVLLGTAPTYVVLAMGWRLLSAFLPARFYQALDDRLCYVQSWLFEFFENYTG 60
QY 61 VOILLGDDLPKKNENIYYLANHSTYDWMYADILAIROMALGHVRYVLEKGLMPLXGC 120
DB 61 VOILLGDDLPKKNENIYYLANHSTYDWMYADILAIROMALGHVRYVLEKGLMPLXGC 120
QY 121 YFAOHGIIYKRSKAKNEKEMRNKLOSVDAGTPMYLVIFPESTRYNEPTKVLASQAF 180
DB 121 YFAOHGIIYKRSKAKNEKEMRNKLOSVDAGTPMYLVIFPESTRYNEPTKVLASQAF 180
QY 181 MAORGLAVLKHVLTPIKATHVAFDCMKNYLDAIYDVTYVEKDGQGRRESPTMEFL 240
DB 181 MAORGLAVLKHVLTPIKATHVAFDCMKNYLDAIYDVTYVEKDGQGRRESPTMEFL 240
QY 241 CKCEPKIHIDRIKDDVPEBOEHMRWLHERFEIKDKMLIEFYSPDERRRKRPFGKS 300
DB 203 CKCEPKIHIDRIKDDVPEBOEHMRWLHERFEIKDKMLIEFYSPDERRRKRPFGKS 262
QY 301 VNSKLSIKKTLPMSMLISGLTAGMLMTDAGRKLVTWYIGTLGCLMWTIKA 353
DB 263 VNSKLSIKKTLPMSMLISGLTAGMLMTDAGRKLVTWYIGTLGCLMWTIKA 315

```

```

RESULT 8
US-09-218-207-134
; Sequence 134, Application US/09218207
; Patent No. 6346381
; GENERAL INFORMATION:
; APPLICANT: Cohen, Daniel
; APPLICANT: Blumenfeld, Marta
; APPLICANT: Ilya, Chumakov
; APPLICANT: Bougueleret, Lydie
; TITLE OF INVENTION: Prostate cancer gene
; FILE REFERENCE: GENSET.018CPI
; CURRENT APPLICATION NUMBER: US/09/218, 207
; CURRENT FILING DATE: 1998-12-22
; EARLIER APPLICATION NUMBER: 08/996,306
; EARLIER FILING DATE: 1997-12-22
; EARLIER APPLICATION NUMBER: 60/099,658
; EARLIER FILING DATE: 1998-09-09
; NUMBER OF SEQ ID NOS: 578
; SOFTWARE: Patent.pm
; SEQ ID NO 134
; LENGTH: 315
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE

```



LOCATION: 81..83  
 OTHER INFORMATION: Box 1  
 FEATURE:  
 NAME/KEY: SITE  
 LOCATION: 160..165  
 OTHER INFORMATION: Box 11  
 US-09-218-207-134

Query Match 87.9%; Score 1644; DB 4; Length 315;  
 Best Local Similarity 89.2%; Pred. No. 6.5e-170; Indels 38; Gaps 1;  
 Matches 315; Conservative 0; Mismatches 0;

QY 1 MRYLLPSVLLGTAFTYVLAMGVMRLSAFLPARFYQALDDRLCYVYOSMVLFFFEENTYG 60  
 DB 1 MRYLLPSVLLGTAFTYVLAMGVMRLSAFLPARFYQALDDRLCYVYOSMVLFFFEENTYG 60  
 QY 61 VOILLYGDLPKKNENIYLANHQSVDWIVADILAIROMALGHVRVYKEGLKMLPLYGC 120  
 DB 61 VOILLYGDLPKKNENIYLANHQSVDWIVADILAIROMALGHVRVYKEGLKMLPLYGC 120  
 QY 121 YFAOHGIVYKRSKAFENKEMRNKLOSVDAGTPMYLVIFPEGSTRYNPQOTVYLSASQAF 180  
 DB 121 YFAOHGIVYKRSKAFENKEMRNKLOSVDAGTPMYLVIFPEGSTRYNPQOTVYLSASQAF 180  
 QY 181 AAORGLAVLKHVLTPIKATHVAFDCMKNYLDIAYDVTVVYEGKDDG-GQRRSPMTTEF 240  
 DB 181 AAOR-----GKDDGQRRSPMTTEF 202  
 QY 241 CKECPKIHIDRIDKKVPEBOEHMRMLHERFEIKDKMLIEFYESPDPERRKRPFGK 300  
 DB 203 CKECPKIHIDRIDKKVPEBOEHMRMLHERFEIKDKMLIEFYESPDPERRKRPFGK 262  
 QY 301 VNSKLSIKKTLPSMLISGLTAGMLMTDAGRKLVTYNTWYIGTLGCLWVTIKA 353  
 DB 263 VNSKLSIKKTLPSMLISGLTAGMLMTDAGRKLVTYNTWYIGTLGCLWVTIKA 315

# RESULT 9

US-09-338-907-74  
 Sequence 74, Application US/0938907  
 Patent No. 6265546  
 GENERAL INFORMATION:  
 APPLICANT: Cohen, Daniel  
 APPLICANT: Blumenfeld, Marta  
 APPLICANT: Ilya, Chumakov  
 APPLICANT: Bougueleret, Lydie  
 TITLE OF INVENTION: PROSTATE CANCER GENE  
 FILE REFERENCE: GENSET.18CPICP  
 CURRENT APPLICATION NUMBER: US/09/338,907  
 EARLIER FILING DATE: 1998-06-23  
 EARLIER APPLICATION NUMBER: 08/996,306  
 EARLIER FILING DATE: 1997-12-22  
 EARLIER APPLICATION NUMBER: 60/099,658  
 EARLIER FILING DATE: 1998-09-09  
 EARLIER APPLICATION NUMBER: 09/218,207  
 EARLIER FILING DATE: 1998-12-22  
 NUMBER OF SEQ ID NOS: 578  
 SOFTWARE: Patent.ppt  
 SEQ ID NO 74  
 LENGTH: 354  
 TYPE: PRT  
 ORGANISM: Mus musculus  
 US-09-338-907-74

Query Match 83.8%; Score 1568.5; DB 4; Length 354;  
 Best Local Similarity 81.4%; Pred. No. 1.2e-161; Indels 1; Gaps 1;  
 Matches 288; Conservative 39; Mismatches 26;

QY 61 VOILLYGDLPKKNENIYLANHQSVDWIVADILAIROMALGHVRVYKEGLKMLPLYGC 120  
 DB 61 VOILLYGDLPKKNENIYLANHQSVDWIVADILAIROMALGHVRVYKEGLKMLPLYGC 120  
 QY 121 YFAOHGIVYKRSKAFENKEMRNKLOSVDAGTPMYLVIFPEGSTRYNPQOTVYLSASQAF 180  
 DB 121 YFAOHGIVYKRSKAFENKEMRNKLOSVDAGTPMYLVIFPEGSTRYNPQOTVYLSASQAF 180  
 QY 181 AAORGLAVLKHVLTPIKATHVAFDCMKNYLDIAYDVTVVYEGKDDG-GQRRSPMTTEF 239  
 DB 181 AAORGLAVLKHVLTPIKATHVAFDCMKNYLDIAYDVTVVYEGKDDG-GQRRSPMTTEF 240  
 QY 240 LCKECPKIHIDRIDKKVPEBOEHMRMLHERFEIKDKMLIEFYESPDPERRKRPFGK 299  
 DB 241 LCKECPKIHIDRIDKKVPEBOEHMRMLHERFEIKDKMLIEFYESPDPERRKRPFGK 300  
 QY 300 VNSKLSIKKTLPSMLISGLTAGMLMTDAGRKLVTYNTWYIGTLGCLWVTIKA 353  
 DB 301 VNSKLSIKKTLPSMLISGLTAGMLMTDAGRKLVTYNTWYIGTLGCLWVTIKA 354

# RESULT 10

US-09-218-207-74  
 Sequence 74, Application US/09218207  
 Patent No. 6346381  
 GENERAL INFORMATION:  
 APPLICANT: Cohen, Daniel  
 APPLICANT: Blumenfeld, Marta  
 APPLICANT: Ilya, Chumakov  
 APPLICANT: Bougueleret, Lydie  
 TITLE OF INVENTION: Prostate cancer gene  
 FILE REFERENCE: GENSET.018CPI  
 CURRENT APPLICATION NUMBER: US/09/218,207  
 EARLIER FILING DATE: 1998-12-22  
 EARLIER APPLICATION NUMBER: 08/996,306  
 EARLIER FILING DATE: 1997-12-22  
 EARLIER APPLICATION NUMBER: 60/099,658  
 EARLIER FILING DATE: 1998-09-09  
 NUMBER OF SEQ ID NOS: 578  
 SOFTWARE: Patent.ppt  
 SEQ ID NO 74  
 LENGTH: 354  
 TYPE: PRT  
 ORGANISM: Mus musculus  
 US-09-218-207-74

Query Match 83.8%; Score 1568.5; DB 4; Length 354;  
 Best Local Similarity 81.4%; Pred. No. 1.2e-161; Indels 1; Gaps 1;  
 Matches 288; Conservative 39; Mismatches 26;

QY 1 MRYLLPSVLLGTAFTYVLAMGVMRLSAFLPARFYQALDDRLCYVYOSMVLFFFEENTYG 60  
 DB 1 MRYLLPSVLLGTAFTYVLAMGVMRLSAFLPARFYQALDDRLCYVYOSMVLFFFEENTYG 60  
 QY 61 VOILLYGDLPKKNENIYLANHQSVDWIVADILAIROMALGHVRVYKEGLKMLPLYGC 120  
 DB 61 VOILLYGDLPKKNENIYLANHQSVDWIVADILAIROMALGHVRVYKEGLKMLPLYGC 120  
 QY 121 YFAOHGIVYKRSKAFENKEMRNKLOSVDAGTPMYLVIFPEGSTRYNPQOTVYLSASQAF 180  
 DB 121 YFAOHGIVYKRSKAFENKEMRNKLOSVDAGTPMYLVIFPEGSTRYNPQOTVYLSASQAF 180  
 QY 181 AAORGLAVLKHVLTPIKATHVAFDCMKNYLDIAYDVTVVYEGKDDG-GQRRSPMTTEF 239  
 DB 181 AAORGLAVLKHVLTPIKATHVAFDCMKNYLDIAYDVTVVYEGKDDG-GQRRSPMTTEF 240  
 QY 240 LCKECPKIHIDRIDKKVPEBOEHMRMLHERFEIKDKMLIEFYESPDPERRKRPFGK 299  
 DB 241 LCKECPKIHIDRIDKKVPEBOEHMRMLHERFEIKDKMLIEFYESPDPERRKRPFGK 300  
 QY 300 VNSKLSIKKTLPSMLISGLTAGMLMTDAGRKLVTYNTWYIGTLGCLWVTIKA 353

Db 301 SVHSRLSVKKTLPSSVLISGLTAVMLMTESGRKLYMGWTLYGLGLMFWIKA 354

## RESULT 11

US-09-338-907-135

; Sequence 135, Application US/09338907

; Patent No. 6265546

; GENERAL INFORMATION:

; APPLICANT: Cohen, Daniel

; APPLICANT: Blumenfeld, Marta

; APPLICANT: Ilya, Chumakov

; APPLICANT: Bougueleret, Lydie

; TITLE OF INVENTION: PROSTATE CANCER GENE

; FILE REFERENCE: GENSET.18CP1CP

; CURRENT APPLICATION NUMBER: US/09/338,907

; EARLIER FILING DATE: 1999-06-23

; EARLIER APPLICATION NUMBER: 08/996,306

; EARLIER FILING DATE: 1997-12-22

; EARLIER APPLICATION NUMBER: 60/099,658

; EARLIER FILING DATE: 1998-09-09

; EARLIER APPLICATION NUMBER: 09/218,207

; EARLIER FILING DATE: 1998-12-22

; NUMBER OF SEQ ID NOS: 578

; SOFTWARE: Patent.pm

; SEQ ID NO 135

; LENGTH: 300

; TYPE: PRT

; ORGANISM: Homo sapiens

; FEATURE:

; NAME/KEY: SITE

; LOCATION: 81..83

; OTHER INFORMATION: Box I

; FEATURE:

; NAME/KEY: SITE

; LOCATION: 160..165

; OTHER INFORMATION: Box II

; US-09-338-907-135

Query Match 83.1%; Score 1555.5; DB 4; Length 300;  
Best Local Similarity 85.0%; Pred. No. 2.4e-160;  
Matches 300; Conservative 0; Mismatches 0; Indels 53; Gaps 1;

QY 1 MRYLPSVVLGTAPTYVVLAMGVWRLLSAFLPARFYQALDDRLCYCYOSWVLEFFENYTG 60  
Db 1 MRYLPSVVLGTAPTYVVLAMGVWRLLSAFLPARFYQALDDRLCYCYOSWVLEFFENYTG 60  
QY 61 VOILLYGDLPKKNENITYLANHOSYVDWIVADILAIROMNLGHVRYVLKGLKWLPLYGC 120  
Db 61 VOILLYGDLPKKNENITYLANHOSYVDWIVADILAIROMNLGHVRYVLKGLKWLPLYGC 120  
QY 121 YFAOHGIIYVKKSAKNEKEMRNKLOSVDAGTPTMLVIFPESTRYNPEOTKVLASQAF 180  
Db 121 YFAOHGIIYVKKSAKNEKEMRNKLOSVDAGTPTMLVIFPESTRYNPEOTKVLASQAF 180  
QY 181 AAORGLAVLKHVLTLPRIKATHVAFDCMKNYLDAIYDVTVYEGKDGCGORRESPTMEFL 240  
Db 181 AAOR-----EFL 187  
QY 241 CKECPRIHIHIDRIDKDVPEOEHRMRWLHERFEIKDKMLIEFYSPDERRRKRPFGKS 300  
Db 188 CKECPRIHIHIDRIDKDVPEOEHRMRWLHERFEIKDKMLIEFYSPDERRRKRPFGKS 247  
QY 301 VNSKLSIKKTLPSSMLISGLTAGMLMTDAGRKLYVNTWIIYGTLLGLMWTIKA 353  
Db 248 VNSKLSIKKTLPSSMLISGLTAGMLMTDAGRKLYVNTWIIYGTLLGLMWTIKA 300

## RESULT 12

US-09-218-207-135

; Sequence 135, Application US/09218207

; Patent No. 6346381

; Patent No. 6346381

; GENERAL INFORMATION:  
; APPLICANT: Cohen, Daniel  
; APPLICANT: Blumenfeld, Marta  
; APPLICANT: Ilya, Chumakov  
; APPLICANT: Bougueleret, Lydie  
; TITLE OF INVENTION: Prostate cancer gene  
; FILE REFERENCE: GENSET.018CP1  
; CURRENT APPLICATION NUMBER: US/09/218,207  
; EARLIER FILING DATE: 1998-12-22  
; EARLIER APPLICATION NUMBER: 08/996,306  
; EARLIER FILING DATE: 1997-12-22  
; EARLIER APPLICATION NUMBER: 60/099,658  
; EARLIER FILING DATE: 1998-09-09  
; NUMBER OF SEQ ID NOS: 578  
; SOFTWARE: Patent.pm  
; SEQ ID NO 135  
; LENGTH: 300  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: SITE  
; LOCATION: 81..83  
; OTHER INFORMATION: Box I  
; FEATURE:  
; NAME/KEY: SITE  
; LOCATION: 160..165  
; OTHER INFORMATION: Box II  
; US-09-218-207-135

Query Match 83.1%; Score 1555.5; DB 4; Length 300;  
Best Local Similarity 85.0%; Pred. No. 2.4e-160;  
Matches 300; Conservative 0; Mismatches 0; Indels 53; Gaps 1;

QY 1 MRYLPSVVLGTAPTYVVLAMGVWRLLSAFLPARFYQALDDRLCYCYOSWVLEFFENYTG 60  
Db 1 MRYLPSVVLGTAPTYVVLAMGVWRLLSAFLPARFYQALDDRLCYCYOSWVLEFFENYTG 60  
QY 61 VOILLYGDLPKKNENITYLANHOSYVDWIVADILAIROMNLGHVRYVLKGLKWLPLYGC 120  
Db 61 VOILLYGDLPKKNENITYLANHOSYVDWIVADILAIROMNLGHVRYVLKGLKWLPLYGC 120  
QY 121 YFAOHGIIYVKKSAKNEKEMRNKLOSVDAGTPTMLVIFPESTRYNPEOTKVLASQAF 180  
Db 121 YFAOHGIIYVKKSAKNEKEMRNKLOSVDAGTPTMLVIFPESTRYNPEOTKVLASQAF 180  
QY 181 AAORGLAVLKHVLTLPRIKATHVAFDCMKNYLDAIYDVTVYEGKDGCGORRESPTMEFL 240  
Db 181 AAOR-----EFL 187  
QY 241 CKECPRIHIHIDRIDKDVPEOEHRMRWLHERFEIKDKMLIEFYSPDERRRKRPFGKS 300  
Db 188 CKECPRIHIHIDRIDKDVPEOEHRMRWLHERFEIKDKMLIEFYSPDERRRKRPFGKS 247  
QY 301 VNSKLSIKKTLPSSMLISGLTAGMLMTDAGRKLYVNTWIIYGTLLGLMWTIKA 353  
Db 248 VNSKLSIKKTLPSSMLISGLTAGMLMTDAGRKLYVNTWIIYGTLLGLMWTIKA 300

## RESULT 13

US-09-338-907-127

; Sequence 127, Application US/09338907

; Patent No. 6265546

; GENERAL INFORMATION:

; APPLICANT: Blumenfeld, Marta

; APPLICANT: Ilya, Chumakov

; APPLICANT: Bougueleret, Lydie

; TITLE OF INVENTION: PROSTATE CANCER GENE

; FILE REFERENCE: GENSET.18CP1CP

; CURRENT APPLICATION NUMBER: US/09/338,907

; EARLIER FILING DATE: 1999-06-23

; EARLIER APPLICATION NUMBER: 08/996,306

```

; EARLIER FILING DATE: 1997-12-22
; EARLIER APPLICATION NUMBER: 60/099,658
; EARLIER FILING DATE: 1998-09-09
; EARLIER APPLICATION NUMBER: 09/218,207
; EARLIER FILING DATE: 1998-12-22
; NUMBER OF SEQ ID NOS: 578
; SOFTWARE: Patent.pm
; SEQ ID NO 127
; LENGTH: 291
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: 98..103
; OTHER INFORMATION: Box II
; FEATURE:
; NAME/KEY: SITE
; LOCATION: 149..157
; OTHER INFORMATION: Box III
US-09-338-907-127
```

```

Query Match      80.1%; Score 1498; DB 4; Length 291;
Best Local Similarity 82.4%; Pred. No. 3,9e-154;
Matches 291; Conservative 0; Mismatches 0; Indels 62; Gaps 1;

QY 1 MRYLLPSVLLGTAPTYVLAMGWRLLSAFLPARFYQALDDRLCYQSVLFFPENYTG 60
    |||||||
DB 1 MRYLLPSVLLGTAPTYVLAMGWRLLSAFLPARFYQALDDRLCYQSVLFFPENYTG 60
    |||||||

QY 61 VOILLYGDLPKKNENIYYLANHOSYDWMIVADILAIRONALGHVRYLKGKMLPLYGC 120
    |
DB 61 V----- 61

QY 121 YFAOHGIIYKRSKAKNEKEMRNKLOSVDAGTPTMLVIFPEGTRYNPEOTKVLASQAF 180
    |||||||
DB 62 ---QHGGIYVKRSKAKNEKEMRNKLOSVDAGTPTMLVIFPEGTRYNPEOTKVLASQAF 118
    |||||||

QY 181 AAOAGLAVLKHYLTPRIKATHVAFDCMKNYLDAIYDVTYVVEGKDDGGQRESPTMTEFL 240
    |||||||
DB 119 AAOAGLAVLKHYLTPRIKATHVAFDCMKNYLDAIYDVTYVVEGKDDGGQRESPTMTEFL 178
    |||||||

QY 241 CKCEPKIHIDRIDKDDVEEOEHMRMLHEFEIKDKMLIEFYSPDERKRRPPGKS 300
    |||||||
DB 179 CKCEPKIHIDRIDKDDVEEOEHMRMLHEFEIKDKMLIEFYSPDERKRRPPGKS 238
    |||||||

QY 301 VNSKLSIKKTLPSMLILSGLTAGMLMTDAGRKLIVNTWIYGTLLGCLMTYIKA 353
    |||||||
DB 239 VNSKLSIKKTLPSMLILSGLTAGMLMTDAGRKLIVNTWIYGTLLGCLMTYIKA 291
    |||||||
```

```

RESULT 14
US-09-218-207-127
; Sequence 127, Application US/09218207
; Patent No. 6346381
; GENERAL INFORMATION:
; APPLICANT: Cohen, Daniel
; APPLICANT: Blumenfeld, Marta
; APPLICANT: Ilya, Chumakov
; APPLICANT: Bougueleret, Lydie
; TITLE OF INVENTION: Prostate cancer gene
; FILE REFERENCE: GENSET.018CPI
; CURRENT APPLICATION NUMBER: US/09/218,207
; CURRENT FILING DATE: 1998-12-22
; EARLIER APPLICATION NUMBER: 08/996,306
; EARLIER FILING DATE: 1997-12-22
; EARLIER APPLICATION NUMBER: 60/099,658
; EARLIER FILING DATE: 1998-09-09
; NUMBER OF SEQ ID NOS: 578
; SOFTWARE: Patent.pm
; SEQ ID NO 127
; LENGTH: 291
; TYPE: PRT
```

```

; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: 98..103
; OTHER INFORMATION: Box II
; FEATURE:
; NAME/KEY: SITE
; LOCATION: 149..157
; OTHER INFORMATION: Box III
US-09-218-207-127
```

```

Query Match      80.1%; Score 1498; DB 4; Length 291;
Best Local Similarity 82.4%; Pred. No. 3,9e-154;
Matches 291; Conservative 0; Mismatches 0; Indels 62; Gaps 1;

QY 1 MRYLLPSVLLGTAPTYVLAMGWRLLSAFLPARFYQALDDRLCYQSVLFFPENYTG 60
    |||||||
DB 1 MRYLLPSVLLGTAPTYVLAMGWRLLSAFLPARFYQALDDRLCYQSVLFFPENYTG 60
    |||||||

QY 61 VOILLYGDLPKKNENIYYLANHOSYDWMIVADILAIRONALGHVRYLKGKMLPLYGC 120
    |
DB 61 V----- 61

QY 121 YFAOHGIIYKRSKAKNEKEMRNKLOSVDAGTPTMLVIFPEGTRYNPEOTKVLASQAF 180
    |||||||
DB 62 ---QHGGIYVKRSKAKNEKEMRNKLOSVDAGTPTMLVIFPEGTRYNPEOTKVLASQAF 118
    |||||||

QY 181 AAOAGLAVLKHYLTPRIKATHVAFDCMKNYLDAIYDVTYVVEGKDDGGQRESPTMTEFL 240
    |||||||
DB 119 AAOAGLAVLKHYLTPRIKATHVAFDCMKNYLDAIYDVTYVVEGKDDGGQRESPTMTEFL 178
    |||||||

QY 241 CKCEPKIHIDRIDKDDVEEOEHMRMLHEFEIKDKMLIEFYSPDERKRRPPGKS 300
    |||||||
DB 179 CKCEPKIHIDRIDKDDVEEOEHMRMLHEFEIKDKMLIEFYSPDERKRRPPGKS 238
    |||||||

QY 301 VNSKLSIKKTLPSMLILSGLTAGMLMTDAGRKLIVNTWIYGTLLGCLMTYIKA 353
    |||||||
DB 239 VNSKLSIKKTLPSMLILSGLTAGMLMTDAGRKLIVNTWIYGTLLGCLMTYIKA 291
    |||||||
```

```

RESULT 15
US-09-338-907-128
; Sequence 128, Application US/09338907
; Patent No. 6265546
; GENERAL INFORMATION:
; APPLICANT: Cohen, Daniel
; APPLICANT: Blumenfeld, Marta
; APPLICANT: Ilya, Chumakov
; APPLICANT: Bougueleret, Lydie
; TITLE OF INVENTION: PROSTATE CANCER GENE
; FILE REFERENCE: GENSET.18CPICP
; CURRENT APPLICATION NUMBER: US/09/338,907
; CURRENT FILING DATE: 1999-06-23
; EARLIER APPLICATION NUMBER: 08/996,306
; EARLIER FILING DATE: 1997-12-22
; EARLIER APPLICATION NUMBER: 60/099,658
; EARLIER FILING DATE: 1998-09-09
; EARLIER APPLICATION NUMBER: 09/218,207
; EARLIER FILING DATE: 1998-12-22
; NUMBER OF SEQ ID NOS: 578
; SOFTWARE: Patent.pm
; SEQ ID NO 128
; LENGTH: 261
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: 68..73
; OTHER INFORMATION: Box II
; FEATURE:
; NAME/KEY: SITE
; LOCATION: 119..127
```

; OTHER INFORMATION: Box III  
US-09-338-907-128

Query Match 70.8%; Score 1325; DB 4; Length 261;  
Best Local Similarity 73.9%; Pred. NO. 1.9e-135;  
Matches 261; Conservative 0; Mismatches 0; Indels 92; Gaps 1;

```
QY 1 MRYLLPSVLLGTAPTYVVLANGVWRLLSAFLPARFYQALDDRLCYQSWVLFEEFENYTG 60
    |||||||
Db 1 MRYLLPSVLLGTAPTYVVLANGVWRLLSAFLPARFYQALDDRLCYQSWVLFEEFENYTG 60
QY 61 VQILLYGDLPKKNENIYYLANHSTVDWIVADILAIRQNALGHVRYVLEKGLKWLPLYGC 120
    ||
Db 61 VQ----- 62
QY 121 YFAQHGGIYVKRSKAFNEKEKRNKLQSYVDAGTTPMYLYIFPEGTRYNPEOTKVLASQAF 180
    |||||||
Db 63 -----MYLYIFPEGTRYNPEOTKVLASQAF 88
QY 181 AAQGLAVLKHHVLPRIKATHVAFDCMKKNYLDATYDVTVVYEGKDDGQRESPTMTEFL 240
    |||||||
Db 89 AAQGLAVLKHHVLPRIKATHVAFDCMKKNYLDATYDVTVVYEGKDDGQRESPTMTEFL 148
QY 241 CKECPKIHIDRIDKDKDVPEEQEHMRRMLHEREIKDKMLIEFYESPDPERRKRFPGKS 300
    |||||||
Db 149 CKECPKIHIDRIDKDKDVPEEQEHMRRMLHEREIKDKMLIEFYESPDPERRKRFPGKS 208
QY 301 VNSKLSIKKTLPSMLISGLTAGMLMTDAGRKLVTNTWYGTGLGCLMVTIKA 353
    |||||||
Db 209 VNSKLSIKKTLPSMLISGLTAGMLMTDAGRKLVTNTWYGTGLGCLMVTIKA 261
```

Search completed: August 28, 2002, 11:08:10  
Job time: 289 sec

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 28, 2002, 11:06:01 ; Search time 306.96 Seconds

(without alignments)  
404.773 Million cell updates/sec

Title: US-09-853-526-4

Sequence: 1 MRYLLPSVLLGAPTYVLA.....YVNTWYTGILGCLWTRKA 353

Scoring table:

BLOSUM62

Searched: 3502263 seqs, 351980561 residues

Total number of hits satisfying chosen parameters: 3502263

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database :

Pending Patents-AA-Main:\*  
1: /cgn2\_6/ptodata/2/paa/PCUS\_COMB.pep:\*  
2: /cgn2\_6/ptodata/2/paa/US08\_COMB.pep:\*  
3: /cgn2\_6/ptodata/2/paa/US07\_COMB.pep:\*  
4: /cgn2\_6/ptodata/2/paa/US080\_COMB.pep:\*  
5: /cgn2\_6/ptodata/2/paa/US081\_COMB.pep:\*  
6: /cgn2\_6/ptodata/2/paa/US082\_COMB.pep:\*  
7: /cgn2\_6/ptodata/2/paa/US083\_COMB.pep:\*  
8: /cgn2\_6/ptodata/2/paa/US084\_COMB.pep:\*  
9: /cgn2\_6/ptodata/2/paa/US085\_COMB.pep:\*  
10: /cgn2\_6/ptodata/2/paa/US086\_COMB.pep:\*  
11: /cgn2\_6/ptodata/2/paa/US087\_COMB.pep:\*  
12: /cgn2\_6/ptodata/2/paa/US088\_COMB.pep:\*  
13: /cgn2\_6/ptodata/2/paa/US089\_COMB.pep:\*  
14: /cgn2\_6/ptodata/2/paa/US090\_COMB.pep:\*  
15: /cgn2\_6/ptodata/2/paa/US091\_COMB.pep:\*  
16: /cgn2\_6/ptodata/2/paa/US092\_COMB.pep:\*  
17: /cgn2\_6/ptodata/2/paa/US093\_COMB.pep:\*  
18: /cgn2\_6/ptodata/2/paa/US094\_COMB.pep:\*  
19: /cgn2\_6/ptodata/2/paa/US095\_COMB.pep:\*  
20: /cgn2\_6/ptodata/2/paa/US096\_COMB.pep:\*  
21: /cgn2\_6/ptodata/2/paa/US097\_COMB.pep:\*  
22: /cgn2\_6/ptodata/2/paa/US098\_COMB.pep:\*  
23: /cgn2\_6/ptodata/2/paa/US099\_COMB.pep:\*  
24: /cgn2\_6/ptodata/2/paa/US100\_COMB.pep:\*  
25: /cgn2\_6/ptodata/2/paa/US101\_COMB.pep:\*  
26: /cgn2\_6/ptodata/2/paa/US60\_COMB.pep:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1871	100.0	353	1	PCT-US01-01431-59
2	1871	100.0	353	1	PCT-US01-01431-79
3	1871	100.0	353	1	PCT-US01-11988-1474
4	1871	100.0	353	1	PCT-US01-11988-1475
5	1871	100.0	353	22	US-09-833-245-1474
6	1871	100.0	353	22	US-09-833-245-1475
7	1871	100.0	353	22	US-09-853-526-4

8	1871	100.0	353	23	US-09-901-484A-4	Sequence 4, Appl
9	1871	100.0	353	23	US-09-915-582-59	Sequence 59, Appl
10	1871	100.0	353	23	US-09-915-582-79	Sequence 79, Appl
11	1871	100.0	353	26	US-60-099-658-4	Sequence 4, Appl
12	1871	100.0	364	18	US-09-436-919-1	Sequence 7, Appl
13	1871	100.0	364	22	US-09-817-910-7	Sequence 1, Appl
14	1871	100.0	364	22	US-09-853-526-5	Sequence 5, Appl
15	1871	100.0	364	23	US-09-901-484A-5	Sequence 5, Appl
16	1871	100.0	364	26	US-60-099-658-5	Sequence 5, Appl
17	1861	99.5	353	18	US-09-488-725A-2736	Sequence 2736, Ap
18	1860	99.4	372	18	US-09-488-725A-6308	Sequence 6308, Ap
19	1644	87.9	315	22	US-09-853-526-134	Sequence 134, App
20	1644	87.9	315	23	US-09-901-484A-134	Sequence 110, App
21	1619.5	86.6	450	26	US-60-245-222-110	Sequence 74, Appl
22	1568.5	83.8	354	22	US-09-853-526-74	Sequence 74, Appl
23	1568.5	83.8	354	23	US-09-901-484A-74	Sequence 74, Appl
24	1568.5	83.8	354	26	US-60-099-658-74	Sequence 309, App
25	1559.5	83.4	1032	26	US-60-212-413-309	Sequence 248, App
26	1559.5	83.4	1032	26	US-60-229-518-248	Sequence 135, App
27	1555.5	83.1	300	22	US-09-853-526-135	Sequence 135, App
28	1555.5	83.1	300	23	US-09-901-484A-135	Sequence 127, App
29	1498	80.1	291	22	US-09-853-526-127	Sequence 127, App
30	1498	80.1	291	23	US-09-901-484A-127	Sequence 47, Appl
31	1427	76.3	269	1	PCT-US01-01327-47	Sequence 128, App
32	1325	70.8	261	22	US-09-853-526-128	Sequence 128, App
33	1325	70.8	261	23	US-09-901-484A-128	Sequence 70, Appl
34	1203	64.3	228	22	US-09-853-526-70	Sequence 70, Appl
35	1203	64.3	228	23	US-09-901-484A-70	Sequence 126, App
36	1203	64.3	228	26	US-60-099-658-70	Sequence 126, App
37	1182.5	63.2	238	22	US-09-853-526-126	Sequence 136, App
38	1182.5	63.2	238	23	US-09-901-484A-126	Sequence 136, App
39	976	52.2	185	22	US-09-853-526-136	Sequence 136, App
40	976	52.2	185	23	US-09-901-484A-136	Sequence 136, App
41	936	50.1	176	18	US-09-436-919-5	Sequence 5, Appl
42	820	43.8	182	22	US-09-853-526-133	Sequence 133, App
43	820	43.8	182	23	US-09-901-484A-133	Sequence 133, App
44	449	24.0	97	22	US-09-853-526-132	Sequence 132, App
45	449	24.0	97	23	US-09-901-484A-132	Sequence 132, App

#### ALIGNMENTS

RESULT 1  
PCT-US01-01431-59  
; Sequence 59, Application PC/TUS0101431  
; GENERAL INFORMATION:  
; APPLICANT: Human Genome Sciences, Inc.  
; TITLE OF INVENTION: 17 human secreted proteins  
; FILE REFERENCE: P5723PCT  
; CURRENT APPLICATION NUMBER: PCT/US01/01431  
; CURRENT FILING DATE: 2001-01-17  
; PRIOR APPLICATION NUMBER: 60/179,065  
; PRIOR FILING DATE: 2000-01-31  
; PRIOR APPLICATION NUMBER: 60/180,628  
; PRIOR FILING DATE: 2000-02-04  
; PRIOR APPLICATION NUMBER: 60/231,968  
; PRIOR FILING DATE: 2000-09-12  
; NUMBER OF SEQ ID NOS: 97  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 59  
; LENGTH: 353  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
PCT-US01-01431-59

Query Match 100.0%; Score 1871; DB 1; Length 353;  
Best Local Similarity 100.0%; Pred. No. 4.1e-184;  
Matches 353; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
1 MRYLLPSVLLGAPTYVLA...GWRLLSAFLPARFOALDDRLCYVQSMVLEFFENYNG 60  
|||||

```
Db 1 MRYLLPSVLLGTAPTYVLAWGVRLSAPLPARFYQALDDRLCYQSVWLEFFENYTG 60
QY 61 VOILLYGDLPKKNENIYLANHSTVDMIVADILAIROMALGHRVYLKEGLKMLPLYGC 120
Db 61 VOILLYGDLPKKNENIYLANHSTVDMIVADILAIROMALGHRVYLKEGLKMLPLYGC 120
QY 121 YFAOHGIIYKRSKAFNEKEKRNKLOSVDAGTPMYLVIPPEGTRYNPEOTKVLASQAF 180
Db 121 YFAOHGIIYKRSKAFNEKEKRNKLOSVDAGTPMYLVIPPEGTRYNPEOTKVLASQAF 180
QY 181 AAOGGLAVLKHLVLPRIKATHVAFDCMKNYLDATYDVTVYVEGDDGQRRSEPTMTTEFL 240
Db 181 AAOGGLAVLKHLVLPRIKATHVAFDCMKNYLDATYDVTVYVEGDDGQRRSEPTMTTEFL 240
QY 241 CKECPKIHIDRIDKDVPEEODEHMRRLHEREIKDKMLIEFYESPDPERRKRPFGKS 300
Db 241 CKECPKIHIDRIDKDVPEEODEHMRRLHEREIKDKMLIEFYESPDPERRKRPFGKS 300
QY 301 VNSKLSIKTKLPSMLILSGLTAGMLMTDAGRKLVTNTWITGTLGCLMTWITKA 353
Db 301 VNSKLSIKTKLPSMLILSGLTAGMLMTDAGRKLVTNTWITGTLGCLMTWITKA 353
```

```
RESULT 2
PCT-US01-01431-79
; Sequence 79, Application PC/TUS0101431
; GENERAL INFORMATION:
; APPLICANT: Human Genome Sciences, Inc.
; TITLE OF INVENTION: 17 human secreted proteins
; FILE REFERENCE: PS723PCT
; CURRENT APPLICATION NUMBER: PCT/US01/01431
; CURRENT FILING DATE: 2001-01-17
; PRIOR APPLICATION NUMBER: 60/179,065
; PRIOR FILING DATE: 2000-01-31
; PRIOR APPLICATION NUMBER: 60/180,628
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: 60/231,968
; PRIOR FILING DATE: 2000-09-12
; NUMBER OF SEQ ID NOS: 97
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 79
; LENGTH: 353
; TYPE: PRT
; ORGANISM: Homo sapiens
PCT-US01-01431-79
```

```
Query Match 100.0%; Score 1871; DB 1; Length 353;
Best Local Similarity 100.0%; Pred. No. 4,1e-184;
Matches 353; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 MRYLLPSVLLGTAPTYVLAWGVRLSAPLPARFYQALDDRLCYQSVWLEFFENYTG 60
Db 1 MRYLLPSVLLGTAPTYVLAWGVRLSAPLPARFYQALDDRLCYQSVWLEFFENYTG 60
QY 61 VOILLYGDLPKKNENIYLANHSTVDMIVADILAIROMALGHRVYLKEGLKMLPLYGC 120
Db 61 VOILLYGDLPKKNENIYLANHSTVDMIVADILAIROMALGHRVYLKEGLKMLPLYGC 120
QY 121 YFAOHGIIYKRSKAFNEKEKRNKLOSVDAGTPMYLVIPPEGTRYNPEOTKVLASQAF 180
Db 121 YFAOHGIIYKRSKAFNEKEKRNKLOSVDAGTPMYLVIPPEGTRYNPEOTKVLASQAF 180
QY 181 AAOGGLAVLKHLVLPRIKATHVAFDCMKNYLDATYDVTVYVEGDDGQRRSEPTMTTEFL 240
Db 181 AAOGGLAVLKHLVLPRIKATHVAFDCMKNYLDATYDVTVYVEGDDGQRRSEPTMTTEFL 240
QY 241 CKECPKIHIDRIDKDVPEEODEHMRRLHEREIKDKMLIEFYESPDPERRKRPFGKS 300
Db 241 CKECPKIHIDRIDKDVPEEODEHMRRLHEREIKDKMLIEFYESPDPERRKRPFGKS 300
QY 301 VNSKLSIKTKLPSMLILSGLTAGMLMTDAGRKLVTNTWITGTLGCLMTWITKA 353
Db 301 VNSKLSIKTKLPSMLILSGLTAGMLMTDAGRKLVTNTWITGTLGCLMTWITKA 353
```

```
Db 301 VNSKLSIKTKLPSMLILSGLTAGMLMTDAGRKLVTNTWITGTLGCLMTWITKA 353
RESULT 3
PCT-US01-11988-1474
; Sequence 1474, Application PC/TUS0111988
; GENERAL INFORMATION:
; APPLICANT: Human Genome Sciences, Inc.
; TITLE OF INVENTION: Albumin Fusion Proteins
; FILE REFERENCE: PF546PCT
; CURRENT APPLICATION NUMBER: PCT/US01/11988
; CURRENT FILING DATE: 2001-01-12
; PRIOR APPLICATION NUMBER: 60/229,358
; PRIOR FILING DATE: 2000-04-12
; PRIOR APPLICATION NUMBER: 60/256,931
; PRIOR FILING DATE: 2000-12-21
; PRIOR APPLICATION NUMBER: 60/199,384
; PRIOR FILING DATE: 2000-04-25
; NUMBER OF SEQ ID NOS: 2267
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1474
; LENGTH: 353
; TYPE: PRT
; ORGANISM: Homo sapiens
PCT-US01-11988-1474
```

```
Query Match 100.0%; Score 1871; DB 1; Length 353;
Best Local Similarity 100.0%; Pred. No. 4,1e-184;
Matches 353; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 MRYLLPSVLLGTAPTYVLAWGVRLSAPLPARFYQALDDRLCYQSVWLEFFENYTG 60
Db 1 MRYLLPSVLLGTAPTYVLAWGVRLSAPLPARFYQALDDRLCYQSVWLEFFENYTG 60
QY 61 VOILLYGDLPKKNENIYLANHSTVDMIVADILAIROMALGHRVYLKEGLKMLPLYGC 120
Db 61 VOILLYGDLPKKNENIYLANHSTVDMIVADILAIROMALGHRVYLKEGLKMLPLYGC 120
QY 121 YFAOHGIIYKRSKAFNEKEKRNKLOSVDAGTPMYLVIPPEGTRYNPEOTKVLASQAF 180
Db 121 YFAOHGIIYKRSKAFNEKEKRNKLOSVDAGTPMYLVIPPEGTRYNPEOTKVLASQAF 180
QY 181 AAOGGLAVLKHLVLPRIKATHVAFDCMKNYLDATYDVTVYVEGDDGQRRSEPTMTTEFL 240
Db 181 AAOGGLAVLKHLVLPRIKATHVAFDCMKNYLDATYDVTVYVEGDDGQRRSEPTMTTEFL 240
QY 241 CKECPKIHIDRIDKDVPEEODEHMRRLHEREIKDKMLIEFYESPDPERRKRPFGKS 300
Db 241 CKECPKIHIDRIDKDVPEEODEHMRRLHEREIKDKMLIEFYESPDPERRKRPFGKS 300
QY 301 VNSKLSIKTKLPSMLILSGLTAGMLMTDAGRKLVTNTWITGTLGCLMTWITKA 353
Db 301 VNSKLSIKTKLPSMLILSGLTAGMLMTDAGRKLVTNTWITGTLGCLMTWITKA 353
```

```
RESULT 4
PCT-US01-11988-1475
; Sequence 1475, Application PC/TUS0111988
; GENERAL INFORMATION:
; APPLICANT: Human Genome Sciences, Inc.
; TITLE OF INVENTION: Albumin Fusion Proteins
; FILE REFERENCE: PF546PCT
; CURRENT APPLICATION NUMBER: PCT/US01/11988
; CURRENT FILING DATE: 2001-01-12
; PRIOR APPLICATION NUMBER: 60/229,358
; PRIOR FILING DATE: 2000-04-12
; PRIOR APPLICATION NUMBER: 60/256,931
; PRIOR FILING DATE: 2000-12-21
; PRIOR APPLICATION NUMBER: 60/199,384
; PRIOR FILING DATE: 2000-04-25
; NUMBER OF SEQ ID NOS: 2267
; SOFTWARE: PatentIn Ver. 2.1
```

; SEQ ID NO 1475  
; LENGTH: 353  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
Pct-US01-11988-1475

Query Match 100.0%; Score 1871; DB 1; Length 353;  
Best Local Similarity 100.0%; Pred. No. 4,1e-184;  
Matches 353; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MRYLPSVLLGTAPTYVLAMGWRLLSAFLPARFYQALDDRLCYQOSWVLEFFENYTG 60  
DB 1 MRYLPSVLLGTAPTYVLAMGWRLLSAFLPARFYQALDDRLCYQOSWVLEFFENYTG 60  
QY 61 VOILLYGDLPKKNENITYLANHSTVDWIVADILAIROMNLGHVRYLKGKGLMPLXGC 120  
DB 61 VOILLYGDLPKKNENITYLANHSTVDWIVADILAIROMNLGHVRYLKGKGLMPLXGC 120  
QY 121 YFAOHGIIYKRSKAKFNEKEMRNKLOSVDAGTPMYLVIPESTRYNPEDTKVLSAQA 180  
DB 121 YFAOHGIIYKRSKAKFNEKEMRNKLOSVDAGTPMYLVIPESTRYNPEDTKVLSAQA 180  
QY 121 YFAOHGIIYKRSKAKFNEKEMRNKLOSVDAGTPMYLVIPESTRYNPEDTKVLSAQA 180  
DB 121 YFAOHGIIYKRSKAKFNEKEMRNKLOSVDAGTPMYLVIPESTRYNPEDTKVLSAQA 180  
QY 181 AAOGLAVLKHVLTLPRIKATHVAFDCMKNYLDIYDVTVYVEGKDDGQRRSPMTTEFL 240  
DB 181 AAOGLAVLKHVLTLPRIKATHVAFDCMKNYLDIYDVTVYVEGKDDGQRRSPMTTEFL 240  
QY 241 CKCEPKIHIIIDRDKKDVPEEOEHMRMLHERFEIKDKMLIEFYESPDERRRKRPFGKS 300  
DB 241 CKCEPKIHIIIDRDKKDVPEEOEHMRMLHERFEIKDKMLIEFYESPDERRRKRPFGKS 300  
QY 301 VNSKLSIKKTLPSMLISGLTAGMLMTDAGRKLVTWVIYGTLLGCLMTWTIKA 353  
DB 301 VNSKLSIKKTLPSMLISGLTAGMLMTDAGRKLVTWVIYGTLLGCLMTWTIKA 353

RESULT 5  
US-09-833-245-1474  
; Sequence 1474, Application US/09833245  
; GENERAL INFORMATION:  
; APPLICANT: Human Genome Sciences, Inc.  
; TITLE OF INVENTION: Albumin Fusion Proteins  
; FILE REFERENCE: PF546PCT  
; CURRENT APPLICATION NUMBER: US/09/833,245  
; CURRENT FILING DATE: 2001-04-12  
; PRIOR APPLICATION NUMBER: 60/229, 358  
; PRIOR FILING DATE: 2000-04-12  
; PRIOR APPLICATION NUMBER: 60/256, 931  
; PRIOR FILING DATE: 2000-12-21  
; PRIOR APPLICATION NUMBER: 60/199, 384  
; PRIOR FILING DATE: 2000-04-25  
; NUMBER OF SEQ ID NOS: 2267  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 1474  
; LENGTH: 353  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-833-245-1474

Query Match 100.0%; Score 1871; DB 22; Length 353;  
Best Local Similarity 100.0%; Pred. No. 4,1e-184;  
Matches 353; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MRYLPSVLLGTAPTYVLAMGWRLLSAFLPARFYQALDDRLCYQOSWVLEFFENYTG 60  
DB 1 MRYLPSVLLGTAPTYVLAMGWRLLSAFLPARFYQALDDRLCYQOSWVLEFFENYTG 60  
QY 61 VOILLYGDLPKKNENITYLANHSTVDWIVADILAIROMNLGHVRYLKGKGLMPLXGC 120  
DB 61 VOILLYGDLPKKNENITYLANHSTVDWIVADILAIROMNLGHVRYLKGKGLMPLXGC 120  
QY 121 YFAOHGIIYKRSKAKFNEKEMRNKLOSVDAGTPMYLVIPESTRYNPEDTKVLSAQA 180

DB 121 YFAOHGIIYKRSKAKFNEKEMRNKLOSVDAGTPMYLVIPESTRYNPEDTKVLSAQA 180  
QY 181 AAOGLAVLKHVLTLPRIKATHVAFDCMKNYLDIYDVTVYVEGKDDGQRRSPMTTEFL 240  
DB 181 AAOGLAVLKHVLTLPRIKATHVAFDCMKNYLDIYDVTVYVEGKDDGQRRSPMTTEFL 240  
QY 241 CKCEPKIHIIIDRDKKDVPEEOEHMRMLHERFEIKDKMLIEFYESPDERRRKRPFGKS 300  
DB 241 CKCEPKIHIIIDRDKKDVPEEOEHMRMLHERFEIKDKMLIEFYESPDERRRKRPFGKS 300  
QY 301 VNSKLSIKKTLPSMLISGLTAGMLMTDAGRKLVTWVIYGTLLGCLMTWTIKA 353  
DB 301 VNSKLSIKKTLPSMLISGLTAGMLMTDAGRKLVTWVIYGTLLGCLMTWTIKA 353

RESULT 6  
US-09-833-245-1475  
; Sequence 1475, Application US/09833245  
; GENERAL INFORMATION:  
; APPLICANT: Human Genome Sciences, Inc.  
; TITLE OF INVENTION: Albumin Fusion Proteins  
; FILE REFERENCE: PF546PCT  
; CURRENT APPLICATION NUMBER: US/09/833,245  
; CURRENT FILING DATE: 2001-04-12  
; PRIOR APPLICATION NUMBER: 60/229, 358  
; PRIOR FILING DATE: 2000-04-12  
; PRIOR APPLICATION NUMBER: 60/256, 931  
; PRIOR FILING DATE: 2000-12-21  
; PRIOR APPLICATION NUMBER: 60/199, 384  
; PRIOR FILING DATE: 2000-04-25  
; NUMBER OF SEQ ID NOS: 2267  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 1475  
; LENGTH: 353  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-833-245-1475

Query Match 100.0%; Score 1871; DB 22; Length 353;  
Best Local Similarity 100.0%; Pred. No. 4,1e-184;  
Matches 353; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MRYLPSVLLGTAPTYVLAMGWRLLSAFLPARFYQALDDRLCYQOSWVLEFFENYTG 60  
DB 1 MRYLPSVLLGTAPTYVLAMGWRLLSAFLPARFYQALDDRLCYQOSWVLEFFENYTG 60  
QY 61 VOILLYGDLPKKNENITYLANHSTVDWIVADILAIROMNLGHVRYLKGKGLMPLXGC 120  
DB 61 VOILLYGDLPKKNENITYLANHSTVDWIVADILAIROMNLGHVRYLKGKGLMPLXGC 120  
QY 121 YFAOHGIIYKRSKAKFNEKEMRNKLOSVDAGTPMYLVIPESTRYNPEDTKVLSAQA 180  
DB 121 YFAOHGIIYKRSKAKFNEKEMRNKLOSVDAGTPMYLVIPESTRYNPEDTKVLSAQA 180  
QY 181 AAOGLAVLKHVLTLPRIKATHVAFDCMKNYLDIYDVTVYVEGKDDGQRRSPMTTEFL 240  
DB 181 AAOGLAVLKHVLTLPRIKATHVAFDCMKNYLDIYDVTVYVEGKDDGQRRSPMTTEFL 240  
QY 241 CKCEPKIHIIIDRDKKDVPEEOEHMRMLHERFEIKDKMLIEFYESPDERRRKRPFGKS 300  
DB 241 CKCEPKIHIIIDRDKKDVPEEOEHMRMLHERFEIKDKMLIEFYESPDERRRKRPFGKS 300  
QY 301 VNSKLSIKKTLPSMLISGLTAGMLMTDAGRKLVTWVIYGTLLGCLMTWTIKA 353  
DB 301 VNSKLSIKKTLPSMLISGLTAGMLMTDAGRKLVTWVIYGTLLGCLMTWTIKA 353

RESULT 7  
US-09-853-526-4  
; Sequence 4, Application US/09853526  
; GENERAL INFORMATION:

APPLICANT: Cohen, Daniel  
APPLICANT: Blumenfeld, Marta  
APPLICANT: Ilya, Chumakov  
TITLE OF INVENTION: PROSTATE CANCER GENE  
FILE REFERENCE: GENSET.18CPICP  
CURRENT APPLICATION NUMBER: US/09/853,526  
PRIORITY FILING DATE: 2001-05-11  
PRIORITY APPLICATION NUMBER: 09/338,907  
PRIORITY FILING DATE: 1999-06-23  
PRIORITY APPLICATION NUMBER: 08/996,306  
PRIORITY FILING DATE: 1997-12-22  
PRIORITY APPLICATION NUMBER: 60/099,658  
PRIORITY FILING DATE: 1998-09-09  
PRIORITY APPLICATION NUMBER: 09/218,207  
NUMBER OF SEQ ID NOS: 578  
SOFTWARE: Patent.pm  
SEQ ID NO 4  
LENGTH: 353  
TYPE: PRT  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: HELIX  
LOCATION: 1..33  
OTHER INFORMATION: Rao and Argos identification method, potential helix  
NAME/KEY: HELIX  
LOCATION: 4..20  
OTHER INFORMATION: Klein, Kanehisa and Delisi identification method, potential  
NAME/KEY: HELIX  
LOCATION: 4..24  
OTHER INFORMATION: Eisenberg, Schwarz, Komarony, Wall identification method, F  
OTHER INFORMATION: helix  
NAME/KEY: MYRISTATE  
LOCATION: 12..16  
OTHER INFORMATION: Prosite match  
NAME/KEY: HELIX  
LOCATION: 50..70  
OTHER INFORMATION: Eisenberg, Schwarz, Komarony, Wall identification method, F  
OTHER INFORMATION: helix  
NAME/KEY: CARBOHYD  
LOCATION: 57..59  
OTHER INFORMATION: Prosite match  
NAME/KEY: HELIX  
LOCATION: 76..96  
OTHER INFORMATION: Eisenberg, Schwarz, Komarony, Wall identification method, F  
OTHER INFORMATION: helix  
NAME/KEY: PHOSPHORYLATION  
LOCATION: 78  
OTHER INFORMATION: potential Tyrosine kinase site, Prosite match  
NAME/KEY: PHOSPHORYLATION  
LOCATION: 84  
OTHER INFORMATION: potential caseine kinase II site, Prosite match  
NAME/KEY: SITE  
LOCATION: 94..115  
OTHER INFORMATION: potential Leucine zipper site, Prosite match  
NAME/KEY: MYRISTATE  
LOCATION: 119..123  
OTHER INFORMATION: potential site, Prosite match  
NAME/KEY: PHOSPHORYLATION  
LOCATION: 133  
OTHER INFORMATION: potential protein kinase C, Prosite match  
NAME/KEY: PHOSPHORYLATION  
LOCATION: 147  
OTHER INFORMATION: potential caseine kinase II site, Prosite match  
NAME/KEY: PHOSPHORYLATION  
LOCATION: 194  
OTHER INFORMATION: potential protein kinase C, Prosite match  
NAME/KEY: PHOSPHORYLATION  
LOCATION: 215  
OTHER INFORMATION: potential Tyrosine kinase site, Prosite match  
NAME/KEY: SULFATATION  
LOCATION: 221

OTHER INFORMATION: Prosite match  
NAME/KEY: PHOSPHORYLATION  
LOCATION: 233  
OTHER INFORMATION: potential cAMP and cGMP dependant protein kinase site,  
OTHER INFORMATION: match  
NAME/KEY: PHOSPHORYLATION  
LOCATION: 235  
OTHER INFORMATION: potential caseine kinase II site, Prosite match  
NAME/KEY: PHOSPHORYLATION  
LOCATION: 306  
OTHER INFORMATION: potential protein kinase C, Prosite match  
NAME/KEY: HELIX  
LOCATION: 310..330  
OTHER INFORMATION: Eisenberg, Schwarz, Komarony, Wall identification method,  
OTHER INFORMATION: helix  
NAME/KEY: MYRISTATE  
LOCATION: 319..323  
OTHER INFORMATION: Prosite match  
NAME/KEY: MYRISTATE  
LOCATION: 323..327  
OTHER INFORMATION: Prosite match  
NAME/KEY: AMIDATION  
LOCATION: 329  
OTHER INFORMATION: Prosite match  
NAME/KEY: HELIX  
LOCATION: 333..353  
OTHER INFORMATION: Eisenberg, Schwarz, Komarony, Wall identification method,  
OTHER INFORMATION: helix  
NAME/KEY: MYRISTATE  
LOCATION: 341..345  
OTHER INFORMATION: Prosite match  
NAME/KEY: PHOSPHORYLATION  
LOCATION: 350  
OTHER INFORMATION: potential protein kinase C, Prosite match  
US-09-853-526-4

Query Match 100.0%; Score 1871; DB 22; Length 353;  
Best Local Similarity 100.0%; Pred. No. 4,1e-184;  
Matches 353; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MRYLLPSVVLGTAFTVYVLMGVWRLSAFLPAFVYALDRLCYQVQSVLTFPEEYTG 60  
DB 1 MRYLLPSVVLGTAFTVYVLMGVWRLSAFLPAFVYALDRLCYQVQSVLTFPEEYTG 60  
QY 61 VOILLYDLPKRNENIYILANHOSTVDVIADILAIROMALGHVRYVLEKGLKMLPLYGC 120  
DB 61 VOILLYDLPKRNENIYILANHOSTVDVIADILAIROMALGHVRYVLEKGLKMLPLYGC 120  
QY 121 YFAOHGGLYKRSKAKFENEKERNKLOSYVDAGTPMYLVIPPEGRVYPEOTKYLSASQAF 180  
DB 121 YFAOHGGLYKRSKAKFENEKERNKLOSYVDAGTPMYLVIPPEGRVYPEOTKYLSASQAF 180  
QY 181 AAORGGLVLRKVLTPRIKATHVAFDCKKNYLDAYDVTYVYEGKDDGGORRESPTMEFL 240  
DB 181 AAORGGLVLRKVLTPRIKATHVAFDCKKNYLDAYDVTYVYEGKDDGGORRESPTMEFL 240  
QY 241 CKECPKIHIDRIIDKDDVEEOEHMRRLHEREIKDKMLIEFYESPDERRRKRPFGKS 300  
DB 241 CKECPKIHIDRIIDKDDVEEOEHMRRLHEREIKDKMLIEFYESPDERRRKRPFGKS 300  
QY 301 VNSKLSIKTKLPMLILSGLTAGMLMTDAGRKLYVNTWYIGTLGCLMVTIKA 353  
DB 301 VNSKLSIKTKLPMLILSGLTAGMLMTDAGRKLYVNTWYIGTLGCLMVTIKA 353

RESULT 8  
US-09-901-484A-4  
; Sequence 4, Application US/09901484A  
; GENERAL INFORMATION:  
; APPLICANT: Cohen, Daniel  
; APPLICANT: Blumenfeld, Marta  
; APPLICANT: Chumakov, Ilya



```

APPLICANT: Bougueleret, Lydie
TITLE OF INVENTION: Prostate Cancer Gene
FILE REFERENCE: GEN-TILXC3D2
CURRENT APPLICATION NUMBER: US/09/901,484A
CURRENT FILING DATE: 2001-07-09
PRIOR APPLICATION NUMBER: US 08/996,306
PRIOR FILING DATE: 1997-12-22
PRIOR APPLICATION NUMBER: US 60/099,658
PRIOR FILING DATE: 1998-09-09
PRIOR APPLICATION NUMBER: US 09/218,207
PRIOR FILING DATE: 1998-12-22
PRIOR APPLICATION NUMBER: US 09/338,907
PRIOR FILING DATE: 1999-06-23
PRIOR APPLICATION NUMBER: US 09/853,526
PRIOR FILING DATE: 2001-05-11
NUMBER OF SEQ ID NOS: 578
SOFTWARE: PatentIn version 3.1
SEQ ID NO 4
LENGTH: 353
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: HELIX
LOCATION: (1)..(33)
OTHER INFORMATION: Rao and Argos identification method, potential helix
NAME/KEY: HELIX
LOCATION: (4)..(20)
OTHER INFORMATION: Klein, Kanehisa and Delisi identification method, potential helix
NAME/KEY: HELIX
LOCATION: (4)..(24)
OTHER INFORMATION: Eisenberg, Schwarz, Komarony, Wall identification method,
OTHER INFORMATION: potential helix
NAME/KEY: LIPID
LOCATION: (12)..(16)
OTHER INFORMATION: MYRISTATE, Prosite match
NAME/KEY: HELIX
LOCATION: (50)..(70)
OTHER INFORMATION: Eisenberg, Schwarz, Komarony, Wall identification method,
OTHER INFORMATION: potential helix
NAME/KEY: CARBOHYD
LOCATION: (57)..(59)
OTHER INFORMATION: Prosite match
NAME/KEY: HELIX
LOCATION: (76)..(96)
OTHER INFORMATION: Eisenberg, Schwarz, Komarony, Wall identification method,
OTHER INFORMATION: potential helix
NAME/KEY: MOD_RES
LOCATION: (78)..(78)
OTHER INFORMATION: PHOSPHORYLATION, potential Tyrosine kinase site, Prosite match
NAME/KEY: MOD_RES
LOCATION: (84)..(84)
OTHER INFORMATION: PHOSPHORYLATION, potential caseine kinase II site, Prosite match
NAME/KEY: SITE
LOCATION: (94)..(115)
OTHER INFORMATION: Potential leucine zipper site, Prosite match
NAME/KEY: LIPID
LOCATION: (119)..(123)
OTHER INFORMATION: MYRISTATE, Prosite match
NAME/KEY: MOD_RES
LOCATION: (133)..(133)
OTHER INFORMATION: PHOSPHORYLATION, potential protein kinase C, Prosite match
NAME/KEY: MOD_RES
LOCATION: (147)..(147)
OTHER INFORMATION: PHOSPHORYLATION, potential caseine kinase II, Prosite match
NAME/KEY: MOD_RES
LOCATION: (194)..(194)
OTHER INFORMATION: PHOSPHORYLATION, potential protein kinase C, Prosite match
NAME/KEY: MOD_RES
LOCATION: (215)..(215)
OTHER INFORMATION: PHOSPHORYLATION, potential Tyrosine kinase site, Prosite match
NAME/KEY: MOD_RES
LOCATION: (221)..(221)
OTHER INFORMATION: SULFATATION, Prosite match

```

```

NAME/KEY: MOD_RES
LOCATION: (233)..(233)
OTHER INFORMATION: PHOSPHORYLATION, potential cAMP and cGMP dependant protein kin
OTHER INFORMATION: site, Prosite match
NAME/KEY: MOD_RES
LOCATION: (235)..(235)
OTHER INFORMATION: PHOSPHORYLATION, potential caseine kinase II site, Prosite mat
NAME/KEY: MOD_RES
LOCATION: (306)..(306)
OTHER INFORMATION: PHOSPHORYLATION, potential protein kinase C, Prosite match
NAME/KEY: HELIX
LOCATION: (310)..(330)
OTHER INFORMATION: Eisenberg, Schwarz, Komarony, Wall identification method,
OTHER INFORMATION: potential helix
NAME/KEY: LIPID
LOCATION: (319)..(323)
OTHER INFORMATION: MYRISTATE, Prosite match
NAME/KEY: LIPID
LOCATION: (323)..(327)
OTHER INFORMATION: MYRISTATE, Prosite match
NAME/KEY: MOD_RES
LOCATION: (329)..(329)
OTHER INFORMATION: AMIDATION, Prosite match
NAME/KEY: HELIX
LOCATION: (333)..(353)
OTHER INFORMATION: Eisenberg, Schwarz, Komarony, Wall identification method,
OTHER INFORMATION: potential helix
NAME/KEY: LIPID
LOCATION: (341)..(345)
OTHER INFORMATION: MYRISTATE, Prosite match
NAME/KEY: MOD_RES
LOCATION: (350)..(350)
OTHER INFORMATION: PHOSPHORYLATION, potential protein kinase C, Prosite match
US-09-901-484A-4

Query Match 100.0%; Score 1871; DB 23; Length 353;
Best Local Similarity 100.0%; Pred. No. 4,1e-184;
Matches 353; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MRYLPSVVLGTAPTYVLAMGWRLLSAFLPARFYQALDDRLCYQSMVLEFFENYTG 60
DB 1 MRYLPSVVLGTAPTYVLAMGWRLLSAFLPARFYQALDDRLCYQSMVLEFFENYTG 60
QY 61 VOILLGDLPRKKNENIYIYANHOSYVDYADILAIRONALGHVRYLKEGKMLPLTGC 120
DB 61 VOILLGDLPRKKNENIYIYANHOSYVDYADILAIRONALGHVRYLKEGKMLPLTGC 120
QY 121 YFAOHGIVYKRSKFNEMKRLQSYDAGTPVLYTFPESTRYNPOTRVLSASQAF 180
DB 121 YFAOHGIVYKRSKFNEMKRLQSYDAGTPVLYTFPESTRYNPOTRVLSASQAF 180
QY 121 YFAOHGIVYKRSKFNEMKRLQSYDAGTPVLYTFPESTRYNPOTRVLSASQAF 180
DB 121 YFAOHGIVYKRSKFNEMKRLQSYDAGTPVLYTFPESTRYNPOTRVLSASQAF 180
QY 181 AAOGLAVLKHVLTFRIRAKTHVAFDCMKNYLDAIYDVTVVYEGKDDGGRRESPTTEPL 240
DB 181 AAOGLAVLKHVLTFRIRAKTHVAFDCMKNYLDAIYDVTVVYEGKDDGGRRESPTTEPL 240
QY 241 CKCEPKIHIDRIDKDDVDEDEHMRMLHEFEIKDKMLIEFYSPPERKRRPGKS 300
DB 241 CKCEPKIHIDRIDKDDVDEDEHMRMLHEFEIKDKMLIEFYSPPERKRRPGKS 300
QY 301 VNSKLSIKKTLPMLILSGITAGMLMTDAGRKLYVTWYIGTLGLMTYTIKA 353
DB 301 VNSKLSIKKTLPMLILSGITAGMLMTDAGRKLYVTWYIGTLGLMTYTIKA 353

RESULT 9
US-09-915-582-59
; Sequence 59, Application US/0915582
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: 17 Human Secreted Proteins
; FILE REFERENCE: PS723P1
; CURRENT APPLICATION NUMBER: US/09/915,582

```

;; CURRENT FILING DATE: 2001-07-27  
;; PRIOR APPLICATION NUMBER: PCT/US01/01431  
;; PRIOR FILING DATE: 2001-01-17  
;; PRIOR APPLICATION NUMBER: 60/179,065  
;; PRIOR FILING DATE: 2000-01-31  
;; PRIOR APPLICATION NUMBER: 60/180,628  
;; PRIOR FILING DATE: 2000-02-04  
;; PRIOR APPLICATION NUMBER: 60/231,968  
;; PRIOR FILING DATE: 2000-09-12  
;; NUMBER OF SEQ ID NOS: 97  
;; SOFTWARE: PatentIn Ver. 2.0  
;; SEQ ID NO: 59  
;; LENGTH: 353  
;; TYPE: PRT  
;; ORGANISM: Homo sapiens  
;; US-09-915-582-59

Query Match 100.0%; Score 1871; DB 23; Length 353;  
Best Local Similarity 100.0%; Pred. No. 4, 1e-184;  
Matches 353; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MRYLLPSVLLGTPATYVLAAGVWRLSAPLPAFYQALDRLCYQSVNLEFFENYTG 60  
DB 1 MRYLLPSVLLGTPATYVLAAGVWRLSAPLPAFYQALDRLCYQSVNLEFFENYTG 60  
QY 61 VOILLYDGLPRKNENIYLANHSTVDMIVADILAIQNALGHVRYLKSGKMLPLYGC 120  
DB 61 VOILLYDGLPRKNENIYLANHSTVDMIVADILAIQNALGHVRYLKSGKMLPLYGC 120  
QY 121 YFAOHGIIYVRSKAKFNEKEKRNKLOSVDAGTPMYLVIFPEGRYVPEQTKVLSAQAF 180  
DB 121 YFAOHGIIYVRSKAKFNEKEKRNKLOSVDAGTPMYLVIFPEGRYVPEQTKVLSAQAF 180  
QY 181 AAOGLAVLKHVLTPIKATHVAFDCKKNYLDAYDVTVYVEGKDDGGORRESPTMEFL 240  
DB 181 AAOGLAVLKHVLTPIKATHVAFDCKKNYLDAYDVTVYVEGKDDGGORRESPTMEFL 240  
QY 241 CKECPKIHIDRIDKKDVEEEOEHMRMWLHERFEIKDKMLIEFYESPDERRRKRPFGKS 300  
DB 241 CKECPKIHIDRIDKKDVEEEOEHMRMWLHERFEIKDKMLIEFYESPDERRRKRPFGKS 300  
QY 301 VNSKLSIKTKLPMSLLISGLTAGMLMTDAGRKLIVNTWIVGTLLGCLMTVITKA 353  
DB 301 VNSKLSIKTKLPMSLLISGLTAGMLMTDAGRKLIVNTWIVGTLLGCLMTVITKA 353

RESULT 10  
US-09-915-582-79  
;; Sequence 79, Application US/09915582  
;; GENERAL INFORMATION:  
;; APPLICANT: Rosen et al.  
;; TITLE OF INVENTION: 17 Human Secreted Proteins  
;; FILE REFERENCE: PS723P1  
;; CURRENT APPLICATION NUMBER: US/09/915,582  
;; PRIOR FILING DATE: 2001-07-27  
;; PRIOR APPLICATION NUMBER: PCT/US01/01431  
;; PRIOR FILING DATE: 2001-01-17  
;; PRIOR APPLICATION NUMBER: 60/179,065  
;; PRIOR FILING DATE: 2000-01-31  
;; PRIOR APPLICATION NUMBER: 60/180,628  
;; PRIOR FILING DATE: 2000-02-04  
;; PRIOR APPLICATION NUMBER: 60/231,968  
;; PRIOR FILING DATE: 2000-09-12  
;; NUMBER OF SEQ ID NOS: 97  
;; SOFTWARE: PatentIn Ver. 2.0  
;; SEQ ID NO: 79  
;; LENGTH: 353  
;; TYPE: PRT  
;; ORGANISM: Homo sapiens  
;; US-09-915-582-79

Query Match 100.0%; Score 1871; DB 23; Length 353;  
Best Local Similarity 100.0%; Pred. No. 4, 1e-184;  
Matches 353; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MRYLLPSVLLGTPATYVLAAGVWRLSAPLPAFYQALDRLCYQSVNLEFFENYTG 60  
DB 1 MRYLLPSVLLGTPATYVLAAGVWRLSAPLPAFYQALDRLCYQSVNLEFFENYTG 60  
QY 61 VOILLYDGLPRKNENIYLANHSTVDMIVADILAIQNALGHVRYLKSGKMLPLYGC 120  
DB 61 VOILLYDGLPRKNENIYLANHSTVDMIVADILAIQNALGHVRYLKSGKMLPLYGC 120  
QY 121 YFAOHGIIYVRSKAKFNEKEKRNKLOSVDAGTPMYLVIFPEGRYVPEQTKVLSAQAF 180  
DB 121 YFAOHGIIYVRSKAKFNEKEKRNKLOSVDAGTPMYLVIFPEGRYVPEQTKVLSAQAF 180  
QY 181 AAOGLAVLKHVLTPIKATHVAFDCKKNYLDAYDVTVYVEGKDDGGORRESPTMEFL 240  
DB 181 AAOGLAVLKHVLTPIKATHVAFDCKKNYLDAYDVTVYVEGKDDGGORRESPTMEFL 240  
QY 241 CKECPKIHIDRIDKKDVEEEOEHMRMWLHERFEIKDKMLIEFYESPDERRRKRPFGKS 300  
DB 241 CKECPKIHIDRIDKKDVEEEOEHMRMWLHERFEIKDKMLIEFYESPDERRRKRPFGKS 300  
QY 301 VNSKLSIKTKLPMSLLISGLTAGMLMTDAGRKLIVNTWIVGTLLGCLMTVITKA 353  
DB 301 VNSKLSIKTKLPMSLLISGLTAGMLMTDAGRKLIVNTWIVGTLLGCLMTVITKA 353

RESULT 11  
US-60-099-658-4  
;; Sequence 4, Application US/60099658  
;; GENERAL INFORMATION:  
;; APPLICANT: Cohen, Daniel  
;; APPLICANT: Chumakov, Ilya  
;; APPLICANT: Blumenfeld, Marta  
;; APPLICANT: Bougenelret, Lydie  
;; TITLE OF INVENTION: Prostate cancer gene  
;; NUMBER OF SEQUENCES: 99  
;; CORRESPONDENCE ADDRESS:  
;; ADDRESSEE: Knobe, Martens, Olson & Bear  
;; STREET: 501 West Broadway  
;; CITY: San Diego  
;; STATE: California  
;; COUNTRY: USA  
;; ZIP: 92101-3505  
;; COMPUTER READABLE FORM:  
;; MEDIUM TYPE: Floppy disk  
;; COMPUTER: IBM PC compatible  
;; OPERATING SYSTEM: Win95  
;; SOFTWARE: Word  
;; CURRENT APPLICATION DATA:  
;; APPLICATION NUMBER: US/60/099,658  
;; FILING DATE:  
;; CLASSIFICATION:  
;; ATTORNEY/AGENT INFORMATION:  
;; NAME: Israelson, Ned A.  
;; REGISTRATION NUMBER: 29,655  
;; REFERENCE/DOCKET NUMBER: GENSET.018APR  
;; TELECOMMUNICATION INFORMATION:  
;; TELEPHONE: (619) 235-8550  
;; TELEFAX: (619) 235-0176  
;; INFORMATION FOR SEQ ID NO: 4:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 353 amino acids  
;; TYPE: AMINO ACID  
;; STRANDEDNESS: SINGLE  
;; TOPOLOGY: LINEAR  
;; MOLECULE TYPE: PROTEIN  
;; ORIGINAL SOURCE:  
;; ORGANISM: Homo sapiens  
;; FEATURE:  
;; NAME/KEY: potential Transmembrane helix

```
LOCATION: 1..33
IDENTIFICATION METHOD: Rao and Argos method
FEATURE:
NAME/KEY: potential Transmembrane helix
LOCATION: 4..20
IDENTIFICATION METHOD: Klein, Kanehisa and Delisi method
FEATURE:
NAME/KEY: potential Transmembrane helix
LOCATION: 4..24
IDENTIFICATION METHOD: Eisenberg, Schwarz, Komarony
and Wall method
FEATURE:
NAME/KEY: potential N-myristoylation site
LOCATION: 12
IDENTIFICATION METHOD: prosite match
FEATURE:
NAME/KEY: potential Transmembrane helix
LOCATION: 50..70
IDENTIFICATION METHOD: Eisenberg, Schwarz, Komarony and Wall method
FEATURE:
NAME/KEY: potential N-glycosylation site
LOCATION: 57
IDENTIFICATION METHOD: prosite match
FEATURE:
NAME/KEY: potential Transmembrane helix
LOCATION: 76..96
IDENTIFICATION METHOD: Eisenberg, Schwarz, Komarony and Wall method
FEATURE:
NAME/KEY: potential Tyrosine kinase phosphorylation site
LOCATION: 78
IDENTIFICATION METHOD: prosite match
FEATURE:
NAME/KEY: potential Casein kinase II phosphorylation site
LOCATION: 84
IDENTIFICATION METHOD: prosite match
FEATURE:
NAME/KEY: potential leucine zipper pattern
LOCATION: 94..115
IDENTIFICATION METHOD: prosite match
FEATURE:
NAME/KEY: potential N-myristoylation site
LOCATION: 119
IDENTIFICATION METHOD: prosite match
FEATURE:
NAME/KEY: potential Protein kinase C phosphorylation site
LOCATION: 133
IDENTIFICATION METHOD: prosite match
FEATURE:
NAME/KEY: potential Casein kinase II phosphorylation site
LOCATION: 147
IDENTIFICATION METHOD: prosite match
FEATURE:
NAME/KEY: potential Protein kinase C phosphorylation site
LOCATION: 194
IDENTIFICATION METHOD: prosite match
FEATURE:
NAME/KEY: potential Tyrosine kinase phosphorylation site
LOCATION: 215
IDENTIFICATION METHOD: prosite match
FEATURE:
NAME/KEY: potential Tyrosine sulfatation site
LOCATION: 221
IDENTIFICATION METHOD: prosite match
FEATURE:
NAME/KEY: potential cAMP- and cGMP-dependent protein kinase phosphorylation site
LOCATION: 233
IDENTIFICATION METHOD: prosite match
FEATURE:
NAME/KEY: potential Casein kinase II phosphorylation site
LOCATION: 235
IDENTIFICATION METHOD: prosite match
FEATURE:
NAME/KEY: potential Protein kinase C phosphorylation site
```

```
LOCATION: 306
IDENTIFICATION METHOD: prosite match
FEATURE:
NAME/KEY: potential Transmembrane helix
LOCATION: 310..330
IDENTIFICATION METHOD: Eisenberg, Schwarz, Komarony and Wall method
FEATURE:
NAME/KEY: potential N-myristoylation site
LOCATION: 319
IDENTIFICATION METHOD: prosite match
FEATURE:
NAME/KEY: potential N-myristoylation site
LOCATION: 323
IDENTIFICATION METHOD: prosite match
FEATURE:
NAME/KEY: potential Amidation site
LOCATION: 329
IDENTIFICATION METHOD: prosite match
FEATURE:
NAME/KEY: potential Transmembrane helix
LOCATION: 333..353
IDENTIFICATION METHOD: Eisenberg, Schwarz, Komarony and Wall method
FEATURE:
NAME/KEY: potential N-myristoylation site
LOCATION: 341
IDENTIFICATION METHOD: prosite match
FEATURE:
NAME/KEY: potential Protein kinase C phosphorylation site
LOCATION: 350
IDENTIFICATION METHOD: prosite match
US-60-099-658-4
```

```
Query Match 100.0% Score 1871; DB: 26; Length 353;
Best Local Similarity 100.0% Pred. No. 4,1e-184;
Matches 353; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 MRYLLPSVVLGTAFTYVLANGVWRLLSAFLPARFYQALDDBLYCVYOSMVLFFPENTG 60
Db 1 MRYLLPSVVLGTAFTYVLANGVWRLLSAFLPARFYQALDDBLYCVYOSMVLFFPENTG 60
Oy 61 VOILLYDLPKNKNIIVLANHSTVDWIVADIIARONALGHVRYVLEGLKMLPLYGC 120
Db 61 VOILLYDLPKNKNIIVLANHSTVDWIVADIIARONALGHVRYVLEGLKMLPLYGC 120
Oy 121 YFAOHGIIYVKSASFENKEMRNKLOSYVDAGTPMVLVIFPEGTRYNEBOTKVLASQAF 180
Db 121 YFAOHGIIYVKSASFENKEMRNKLOSYVDAGTPMVLVIFPEGTRYNEBOTKVLASQAF 180
Oy 181 AAORGLAVLKHVLPRIKATVAFPCMKNYIDAIVDVTVVEGKDDGQRRSPMTTEFL 240
Db 181 AAORGLAVLKHVLPRIKATVAFPCMKNYIDAIVDVTVVEGKDDGQRRSPMTTEFL 240
Oy 241 CKCPKIHIDRIDKDVPEQEHMRRLHREFIKDKMLIEFESDPERRKRPFGKS 300
Db 241 CKCPKIHIDRIDKDVPEQEHMRRLHREFIKDKMLIEFESDPERRKRPFGKS 300
Oy 301 VNSKSTIKTLPMSLILSGLTAGMLMTDAGRKLYVNTVIYGLLGLCLWTTIKA 353
Db 301 VNSKSTIKTLPMSLILSGLTAGMLMTDAGRKLYVNTVIYGLLGLCLWTTIKA 353
```

```
RESULT 12
US-09-436-919-1
Sequence 1, Application US/09436919A
GENERAL INFORMATION:
APPLICANT: Leung, David W
TITLE OF INVENTION: Human Lysophosphatidic Acid Acyltransferase-epsilon
FILE REFERENCE: 1801B
CURRENT APPLICATION NUMBER: US/09/436,919A
CURRENT FILING DATE: 1999-11-09
NUMBER OF SEQ ID NOS: 5
SOFTWARE: PatentIn Ver. 2.0
```

```
; SEQ ID NO 1
; LENGTH: 364
; TYPE: PRF
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: hLPAAteps11on
US-09-436-919-1
```

```
Query Match          100.0%; Score 1871; DB 18; Length 364;
Best Local Similarity 100.0%; Pred. No. 4.2e-184;
Matches 353; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 MRYLLPSVVLGTAPTYVLAMGWRLLSAFLPARFYQALDDRLCYCYOSMVLFFEFENTYG 60
    |||||||
DB 12 MRYLLPSVVLGTAPTYVLAMGWRLLSAFLPARFYQALDDRLCYCYOSMVLFFEFENTYG 71
    |||||||
QY 61 VOILLYGDLPKKNKENTIIYLANHOSTYDVIADILAIROMALGHVRVYVLEKGLKWLPLYGC 120
    |||||||
DB 72 VOILLYGDLPKKNKENTIIYLANHOSTYDVIADILAIROMALGHVRVYVLEKGLKWLPLYGC 131
    |||||||
QY 121 YFAOHGIIYVKSARKFNEKEMRNKLOSVDAGTPMYLVIFPEGSTRYNPEQTKVLSASQAF 180
    |||||||
DB 132 YFAOHGIIYVKSARKFNEKEMRNKLOSVDAGTPMYLVIFPEGSTRYNPEQTKVLSASQAF 191
    |||||||
QY 181 AAOGLAVLKHVLTTPRIKATHVAFDCMKNYLDAIYDVTVVYEGKDDGGRRESPTMTTEFL 240
    |||||||
DB 192 AAOGLAVLKHVLTTPRIKATHVAFDCMKNYLDAIYDVTVVYEGKDDGGRRESPTMTTEFL 251
    |||||||
QY 241 CKECPKIHIDRIDKDVPEOEHRMRWLHERFEIKDKMLIEFYESPDPERRKRPFGKS 300
    |||||||
DB 252 CKECPKIHIDRIDKDVPEOEHRMRWLHERFEIKDKMLIEFYESPDPERRKRPFGKS 311
    |||||||
QY 301 VNSKLSIKKTLPSMLILSGLTAGMLMTDAGRKLVTWMTWYIGTLGLCMTWTJKA 353
    |||||||
DB 312 VNSKLSIKKTLPSMLILSGLTAGMLMTDAGRKLVTWMTWYIGTLGLCMTWTJKA 364
    |||||||
```

## RESULT 13

```
; US-09-817-910-7
; Sequence 7, Application US/09817910
; GENERAL INFORMATION:
; APPLICANT: Meyers, Rachel A.
; APPLICANT: Macbeth, Kyle J.
; APPLICANT: Williamson, Mark
; APPLICANT: Rudolph-Owen, Laura A.
; APPLICANT: Tsai, Fong-Ying
; TITLE OF INVENTION: 46743 and 27417 NOVEL HUMAN
; FILE REFERENCE: 10448-032001
; CURRENT FILING DATE: US/09/817,910
; PRIOR APPLICATION NUMBER: 2001-03-26
; PRIOR FILING DATE: 2000-03-24
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 7
; LENGTH: 364
; TYPE: PRF
; ORGANISM: Homo sapiens
US-09-817-910-7
```

```
Query Match          100.0%; Score 1871; DB 22; Length 364;
Best Local Similarity 100.0%; Pred. No. 4.2e-184;
Matches 353; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 MRYLLPSVVLGTAPTYVLAMGWRLLSAFLPARFYQALDDRLCYCYOSMVLFFEFENTYG 60
    |||||||
DB 12 MRYLLPSVVLGTAPTYVLAMGWRLLSAFLPARFYQALDDRLCYCYOSMVLFFEFENTYG 71
    |||||||
QY 61 VOILLYGDLPKKNKENTIIYLANHOSTYDVIADILAIROMALGHVRVYVLEKGLKWLPLYGC 120
    |||||||
```

```
DB 72 VOILLYGDLPKKNKENTIIYLANHOSTYDVIADILAIROMALGHVRVYVLEKGLKWLPLYGC 131
    |||||||
QY 121 YFAOHGIIYVKSARKFNEKEMRNKLOSVDAGTPMYLVIFPEGSTRYNPEQTKVLSASQAF 180
    |||||||
DB 132 YFAOHGIIYVKSARKFNEKEMRNKLOSVDAGTPMYLVIFPEGSTRYNPEQTKVLSASQAF 191
    |||||||
QY 181 AAOGLAVLKHVLTTPRIKATHVAFDCMKNYLDAIYDVTVVYEGKDDGGRRESPTMTTEFL 240
    |||||||
DB 192 AAOGLAVLKHVLTTPRIKATHVAFDCMKNYLDAIYDVTVVYEGKDDGGRRESPTMTTEFL 251
    |||||||
QY 241 CKECPKIHIDRIDKDVPEOEHRMRWLHERFEIKDKMLIEFYESPDPERRKRPFGKS 300
    |||||||
DB 252 CKECPKIHIDRIDKDVPEOEHRMRWLHERFEIKDKMLIEFYESPDPERRKRPFGKS 311
    |||||||
QY 301 VNSKLSIKKTLPSMLILSGLTAGMLMTDAGRKLVTWMTWYIGTLGLCMTWTJKA 353
    |||||||
DB 312 VNSKLSIKKTLPSMLILSGLTAGMLMTDAGRKLVTWMTWYIGTLGLCMTWTJKA 364
    |||||||
```

## RESULT 14

```
; US-09-853-526-5
; Sequence 5, Application US/09853526
; GENERAL INFORMATION:
; APPLICANT: Blumenfeld, Marta
; APPLICANT: Ilyia, Chumakov
; APPLICANT: Bougueleret, Lydie
; TITLE OF INVENTION: PROSTATE CANCER GENE
; FILE REFERENCE: GENSET.18CPLCP
; CURRENT APPLICATION NUMBER: US/09/853,526
; PRIOR FILING DATE: 2001-05-11
; PRIOR APPLICATION NUMBER: 09/338,907
; PRIOR FILING DATE: 1999-06-23
; PRIOR APPLICATION NUMBER: 08/996,306
; PRIOR FILING DATE: 1997-12-22
; PRIOR APPLICATION NUMBER: 60/099,658
; PRIOR FILING DATE: 1998-09-09
; PRIOR APPLICATION NUMBER: 09/218,207
; PRIOR FILING DATE: 1998-12-22
; NUMBER OF SEQ ID NOS: 578
; SOFTWARE: Patent.pm
; SEQ ID NO 5
; LENGTH: 364
; TYPE: PRF
; ORGANISM: Homo sapiens
US-09-853-526-5
```

```
Query Match          100.0%; Score 1871; DB 22; Length 364;
Best Local Similarity 100.0%; Pred. No. 4.2e-184;
Matches 353; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 MRYLLPSVVLGTAPTYVLAMGWRLLSAFLPARFYQALDDRLCYCYOSMVLFFEFENTYG 60
    |||||||
DB 12 MRYLLPSVVLGTAPTYVLAMGWRLLSAFLPARFYQALDDRLCYCYOSMVLFFEFENTYG 71
    |||||||
QY 61 VOILLYGDLPKKNKENTIIYLANHOSTYDVIADILAIROMALGHVRVYVLEKGLKWLPLYGC 120
    |||||||
DB 72 VOILLYGDLPKKNKENTIIYLANHOSTYDVIADILAIROMALGHVRVYVLEKGLKWLPLYGC 131
    |||||||
QY 121 YFAOHGIIYVKSARKFNEKEMRNKLOSVDAGTPMYLVIFPEGSTRYNPEQTKVLSASQAF 180
    |||||||
DB 132 YFAOHGIIYVKSARKFNEKEMRNKLOSVDAGTPMYLVIFPEGSTRYNPEQTKVLSASQAF 191
    |||||||
QY 181 AAOGLAVLKHVLTTPRIKATHVAFDCMKNYLDAIYDVTVVYEGKDDGGRRESPTMTTEFL 240
    |||||||
DB 192 AAOGLAVLKHVLTTPRIKATHVAFDCMKNYLDAIYDVTVVYEGKDDGGRRESPTMTTEFL 251
    |||||||
QY 241 CKECPKIHIDRIDKDVPEOEHRMRWLHERFEIKDKMLIEFYESPDPERRKRPFGKS 300
    |||||||
DB 252 CKECPKIHIDRIDKDVPEOEHRMRWLHERFEIKDKMLIEFYESPDPERRKRPFGKS 311
    |||||||
QY 301 VNSKLSIKKTLPSMLILSGLTAGMLMTDAGRKLVTWMTWYIGTLGLCMTWTJKA 353
    |||||||
```

Db 312 VNSKLSIKKTLPSMLILSGLTAGMLMTDAGRKLVTWTVYGTLLGCLWVTIKA 364

RESULT 15  
US-09-901-484A-5  
; Sequence 5, Application US/09901484A  
; GENERAL INFORMATION:  
; APPLICANT: Cohen, Daniel  
; APPLICANT: Blumenfeld, Marta  
; APPLICANT: Chumakov, Ilya  
; APPLICANT: Bouquelere, Lydie  
; TITLE OF INVENTION: Prostate Cancer Gene  
; FILE REFERENCE: GEN-T11XC3D2  
; CURRENT APPLICATION NUMBER: US/09/901,484A  
; CURRENT FILING DATE: 2001-07-09  
; PRIOR APPLICATION NUMBER: US 08/996,306  
; PRIOR FILING DATE: 1997-12-22  
; PRIOR APPLICATION NUMBER: US 60/099,658  
; PRIOR FILING DATE: 1998-09-09  
; PRIOR APPLICATION NUMBER: US 09/218,207  
; PRIOR FILING DATE: 1998-12-22  
; PRIOR APPLICATION NUMBER: US 09/338,907  
; PRIOR FILING DATE: 1999-06-23  
; PRIOR APPLICATION NUMBER: US 09/853,526  
; PRIOR FILING DATE: 2001-05-11  
; NUMBER OF SEQ ID NOS: 578  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 5  
; LENGTH: 364  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-901-484A-5

Query Match 100.0%; Score 1871; DB 23; Length 364;  
Best Local Similarity 100.0%; Pred. No. 4,2e-184;  
Matches 353; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MRYLLPSVVLGTAPTYVLAMGVMRLSFLPARFYQALDDRLCYCYOSMVLFFPENTYG 60  
Db 12 MRYLLPSVVLGTAPTYVLAMGVMRLSFLPARFYQALDDRLCYCYOSMVLFFPENTYG 71  
Qy 61 VOILLYGLDPKKNENIYLANHQSIVDMIVADILAIROMALGHVRYLKEGLKMLPLYGC 120  
Db 72 VOILLYGLDPKKNENIYLANHQSIVDMIVADILAIROMALGHVRYLKEGLKMLPLYGC 131  
Qy 121 YFAQHGGIYVKRSKAFNEKEMRNKLQSYVDAGTPMYLVIFPEGTRYNPQOTVLSASQAF 180  
Db 132 YFAQHGGIYVKRSKAFNEKEMRNKLQSYVDAGTPMYLVIFPEGTRYNPQOTVLSASQAF 191  
Qy 181 AAOBGLAVLKHLTPRIKATHAFDCMKNYLDAIVDVTVVYEGKDDGGRRESPTTEFL 240  
Db 192 AAOBGLAVLKHLTPRIKATHAFDCMKNYLDAIVDVTVVYEGKDDGGRRESPTTEFL 251  
Qy 241 CKCEPKIHIDRIKDVPEEQEHMRRLHERFEIKDKMLIEFYESPDERRRKRPFGKS 300  
Db 252 CKCEPKIHIDRIKDVPEEQEHMRRLHERFEIKDKMLIEFYESPDERRRKRPFGKS 311  
Qy 301 VNSKLSIKKTLPSMLILSGLTAGMLMTDAGRKLVTWTVYGTLLGCLWVTIKA 353  
Db 312 VNSKLSIKKTLPSMLILSGLTAGMLMTDAGRKLVTWTVYGTLLGCLWVTIKA 364

Search completed: August 28, 2002, 11:14:10  
Job time: 489 sec



GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 28, 2002, 11:06:26 ; Search time 84.08 Seconds

(without alignments)  
1016.222 Million cell updates/sec

Title: US-09-853-526-4

Perfect score: 1871

Sequence: 1 MRYLPSVVLGTAPTYVLA.....YVNTWLYGTLGLMTYIKR 353

Scoring table:

BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 747981 seqs, 242050750 residues

Total number of hits satisfying chosen parameters: 747981

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 100%

Listing first 45 summaries

Database : Pending\_Patents\_AA\_New:\*  
1: /cgn2\_6/ptodata/1/paa/PCr\_NEW\_COMB.pep:\*  
2: /cgn2\_6/ptodata/1/paa/US06\_NEW\_COMB.pep:\*  
3: /cgn2\_6/ptodata/1/paa/US07\_NEW\_COMB.pep:\*  
4: /cgn2\_6/ptodata/1/paa/US08\_NEW\_COMB.pep:\*  
5: /cgn2\_6/ptodata/1/paa/US09\_NEW\_COMB.pep:\*  
6: /cgn2\_6/ptodata/1/paa/US10\_NEW\_COMB.pep:\*  
7: /cgn2\_6/ptodata/1/paa/US60\_NEW\_COMB.pep:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1871	100.0	364	US-10-184-648-63	Sequence 63, Appl
2	1868	99.8	353	US-09-629-469A-13028	Sequence 13028, A
3	1427	76.3	269	US-10-074-045-47	Sequence 47, Appl
4	390.5	20.9	375	US-09-935-625-11061	Sequence 11061, A
5	390.5	20.9	375	US-09-935-625-14197	Sequence 14197, A
6	390.5	20.9	375	US-09-935-625-14201	Sequence 14201, A
7	390.5	20.9	375	US-09-935-625-14254	Sequence 14254, A
8	390.5	20.9	375	US-09-935-625-23009	Sequence 23009, A
9	390.5	20.9	375	US-09-935-625-23025	Sequence 23025, A
10	387.5	20.7	351	US-09-935-625-11062	Sequence 11062, A
11	387.5	20.7	351	US-09-935-625-14198	Sequence 14198, A
12	387.5	20.7	351	US-09-935-625-14202	Sequence 14202, A
13	387.5	20.7	351	US-09-935-625-14255	Sequence 14255, A
14	387.5	20.7	351	US-09-935-625-23010	Sequence 23010, A
15	387.5	20.7	351	US-09-935-625-23026	Sequence 23026, A
16	387.5	20.7	375	US-09-935-625-11471	Sequence 11471, A
17	384.5	20.6	351	US-09-935-625-11472	Sequence 11472, A
18	374.5	20.0	343	US-09-935-625-11063	Sequence 11063, A
19	374.5	20.0	343	US-09-935-625-11473	Sequence 11473, A
20	374.5	20.0	343	US-09-935-625-14199	Sequence 14199, A
21	374.5	20.0	343	US-09-935-625-14203	Sequence 14203, A
22	374.5	20.0	343	US-09-935-625-14256	Sequence 14256, A
23	374.5	20.0	343	US-09-935-625-23011	Sequence 23011, A
24	374.5	20.0	343	US-09-935-625-23027	Sequence 23027, A
25	371.5	19.9	400	US-60-391-781-1157	Sequence 1157, Ap
26	305.5	16.3	414	US-10-121-062-314	Sequence 314, App

27	305.5	16.3	414	US-10-063-502-102	Sequence 102, App
28	305.5	16.3	414	US-10-063-510-102	Sequence 102, App
29	305.5	16.3	414	US-10-063-512-102	Sequence 102, App
30	305.5	16.3	414	US-10-063-513-102	Sequence 102, App
31	305.5	16.3	414	US-10-063-514-102	Sequence 102, App
32	305.5	16.3	414	US-10-063-515-102	Sequence 102, App
33	305.5	16.3	414	US-10-063-516-102	Sequence 102, App
34	305.5	16.3	414	US-10-063-517-102	Sequence 102, App
35	305.5	16.3	414	US-10-063-518-102	Sequence 102, App
36	305.5	16.3	414	US-10-063-519-102	Sequence 102, App
37	305.5	16.3	414	US-10-063-520-102	Sequence 102, App
38	305.5	16.3	414	US-10-063-521-102	Sequence 102, App
39	305.5	16.3	414	US-10-063-523-102	Sequence 102, App
40	305.5	16.3	414	US-10-063-524-102	Sequence 102, App
41	305.5	16.3	414	US-10-063-525-102	Sequence 102, App
42	305.5	16.3	414	US-10-063-526-102	Sequence 102, App
43	305.5	16.3	414	US-10-063-527-102	Sequence 102, App
44	305.5	16.3	414	US-10-063-528-102	Sequence 102, App
45	305.5	16.3	414	US-10-063-529-102	Sequence 102, App

## ALIGNMENTS

RESULT 1  
US-10-184-648-63  
Sequence 63, Application US/10184648  
GENERAL INFORMATION:  
APPLICANT: Meyers, Rachel E.  
APPLICANT: Williamson, Mark  
APPLICANT: Tsai, Fong-Ying  
APPLICANT: Hunter, John Y.  
APPLICANT: Macbeth, Kyle J.  
APPLICANT: Rudolph-Owen, Laura A.  
APPLICANT: Leibly, Kevin R.  
APPLICANT: Kapeller-Libermann, Rosana  
APPLICANT: Olandt, Peter J.  
TITLE OF INVENTION: NOVEL HUMAN TRANSFERASE FAMILY MEMBERS AND USES THEREOF  
FILE REFERENCE: 10448-192001  
CURRENT APPLICATION NUMBER: US/10/184,648  
CURRENT FILING DATE: 2002-06-27  
PRIOR APPLICATION NUMBER: US 09/815,028  
PRIOR FILING DATE: 2001-03-22  
PRIOR APPLICATION NUMBER: PCT/US01/09338  
PRIOR FILING DATE: 2001-03-22  
PRIOR APPLICATION NUMBER: US 60/191,964  
PRIOR FILING DATE: 2000-03-24  
PRIOR APPLICATION NUMBER: US 09/801,220  
PRIOR FILING DATE: 2001-03-07  
PRIOR APPLICATION NUMBER: PCT/US01/07289  
PRIOR FILING DATE: 2001-03-07  
PRIOR APPLICATION NUMBER: US 60/187,456  
PRIOR FILING DATE: 2000-03-07  
PRIOR APPLICATION NUMBER: US 09/816,714  
PRIOR FILING DATE: 2001-03-23  
PRIOR APPLICATION NUMBER: PCT/US01/09468  
PRIOR FILING DATE: 2001-03-23  
PRIOR APPLICATION NUMBER: US 60/191,865  
PRIOR FILING DATE: 2000-03-24  
PRIOR APPLICATION NUMBER: US 09/844,948  
PRIOR FILING DATE: 2001-04-27  
PRIOR APPLICATION NUMBER: PCT/US01/13805  
PRIOR FILING DATE: 2001-04-27  
PRIOR APPLICATION NUMBER: US 60/200,604  
PRIOR FILING DATE: 2000-04-28  
PRIOR APPLICATION NUMBER: US 09/861,164  
PRIOR FILING DATE: 2001-05-18  
PRIOR APPLICATION NUMBER: PCT/US01/16232  
PRIOR FILING DATE: 2001-05-18  
PRIOR APPLICATION NUMBER: US 60/205,408  
PRIOR FILING DATE: 2000-05-19  
PRIOR APPLICATION NUMBER: US 09/883,060  
PRIOR FILING DATE: 2001-06-15

```

; PRIOR APPLICATION NUMBER: PCT/US01/19138
; PRIOR FILING DATE: 2001-06-15
; PRIOR APPLICATION NUMBER: US 60/212, 079
; PRIOR FILING DATE: 2000-06-15
; PRIOR APPLICATION NUMBER: US 09/962, 678
; PRIOR FILING DATE: 2001-09-25
; PRIOR APPLICATION NUMBER: PCT/US01/29963
; PRIOR FILING DATE: 2001-09-25
; PRIOR APPLICATION NUMBER: US 60/235, 044
; PRIOR FILING DATE: 2000-09-25
; PRIOR APPLICATION NUMBER: US 09/973, 457
; PRIOR FILING DATE: 2001-10-09
; PRIOR APPLICATION NUMBER: US 60/238, 849
; PRIOR FILING DATE: 2000-10-06
; PRIOR APPLICATION NUMBER: US 10/072, 285
; PRIOR FILING DATE: 2002-02-08
; PRIOR APPLICATION NUMBER: PCT/US02/03736
; PRIOR FILING DATE: 2002-02-08
; PRIOR APPLICATION NUMBER: US 60/267, 494
; PRIOR FILING DATE: 2001-02-08
; PRIOR APPLICATION NUMBER: US 09/817, 910
; PRIOR FILING DATE: 2001-03-26
; PRIOR APPLICATION NUMBER: PCT/US01/09633
; PRIOR FILING DATE: 2001-03-26
; PRIOR APPLICATION NUMBER: US 60/192, 092
; PRIOR FILING DATE: 2000-03-24
; PRIOR APPLICATION NUMBER: US 09/842, 528
; PRIOR FILING DATE: 2001-04-25
; PRIOR APPLICATION NUMBER: PCT/US01/40607
; PRIOR FILING DATE: 2001-04-25
; PRIOR APPLICATION NUMBER: US 60/199, 500
; PRIOR FILING DATE: 2000-04-25
; PRIOR APPLICATION NUMBER: US 09/882, 836
; PRIOR FILING DATE: 2001-06-15
; PRIOR APPLICATION NUMBER: PCT/US01/19543
; PRIOR FILING DATE: 2001-06-15
; PRIOR APPLICATION NUMBER: US 60/211, 730
; PRIOR FILING DATE: 2000-06-15
; PRIOR APPLICATION NUMBER: US 09/882, 872
; PRIOR FILING DATE: 2001-06-15
; PRIOR APPLICATION NUMBER: PCT/US01/19153
; PRIOR FILING DATE: 2001-06-15
; PRIOR APPLICATION NUMBER: US 60/212, 077
; PRIOR FILING DATE: 2000-06-15
; NUMBER OF SEQ ID NOS: 90
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 63
; LENGTH: 364
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-184-648-63
```

```

Query Match          100.0%; Score 1871; DB 6; Length 364;
Best Local Similarity 100.0%; Pred. No. 3.7e-167;
Matches 353; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MRYLPSVVLGTAPTYVLAMGWRLLSAFLPARFYQALDRLRYCYQSMVLEFFENYTG 60
    |||||||
DB 12 MRYLPSVVLGTAPTYVLAMGWRLLSAFLPARFYQALDRLRYCYQSMVLEFFENYTG 71
    |||||||
QY 61 VOILLYGLDLPKKNENITYLANHOSYVDWIVADILAIROMALGHVRVYVLEKGLKMLPLYGC 120
    |||||||
DB 72 VOILLYGLDLPKKNENITYLANHOSYVDWIVADILAIROMALGHVRVYVLEKGLKMLPLYGC 131
    |||||||
QY 121 YFAOHGCIYVKSASAKFNEKEMRNKLOSYYVDAGTAPMTLVIFPEGTRYNPQOTVLSASQAF 180
    |||||||
DB 132 YFAOHGCIYVKSASAKFNEKEMRNKLOSYYVDAGTAPMTLVIFPEGTRYNPQOTVLSASQAF 191
    |||||||
QY 181 AAORGLAVLKHVLTPIRIKATHVAFDCMKNYLDAIYDVTVVYEGKDGGRRESPTMTFEL 240
    |||||||
DB 192 AAORGLAVLKHVLTPIRIKATHVAFDCMKNYLDAIYDVTVVYEGKDGGRRESPTMTFEL 251
    |||||||
```

```

QY 241 CKECPKIHIDRIDKDVPEOEHRMRWLHERFEIKDKMLIEFYSPDERKKRPPGKS 300
    |||||||
DB 252 CKECPKIHIDRIDKDVPEOEHRMRWLHERFEIKDKMLIEFYSPDERKKRPPGKS 311
    |||||||
QY 301 VNSKLSIKKTPSMLLSGLTGMLTGDKRKLYVTWYTGTLGGLWMTIRA 353
    |||||||
DB 312 VNSKLSIKKTPSMLLSGLTGMLTGDKRKLYVTWYTGTLGGLWMTIRA 364
    |||||||
```

## RESULT 2

US-09-629-469A-13028

; Sequence 13028, Application US/09629469A

; GENERAL INFORMATION:

; APPLICANT: OTA, TOSHIO

; APPLICANT: ISOGAI, TAKAO

; APPLICANT: NISHIKAWA, TETSUO

; APPLICANT: HAYASHI, KOJI

; APPLICANT: SAITO, KAORU

; APPLICANT: YAMAMOTO, JUNICHI

; APPLICANT: ISHII, SHIZUKO

; APPLICANT: SUGIYAMA, TOMOYASU

; APPLICANT: WAKAMATSU, AI

; APPLICANT: NAGAI, KEIICHI

; APPLICANT: OTSUKI, TETSUJI

; TITLE OF INVENTION: PRIMERS FOR SYNTHESIZING FULL-LENGTH cDNA AND THEIR USE

; FILE REFERENCE: 084335/0123

; CURRENT APPLICATION NUMBER: US/09/629,469A

; CURRENT FILING DATE: 2000-07-28

; PRIOR APPLICATION NUMBER: JP 1999-248036

; PRIOR FILING DATE: 1999-07-29

; PRIOR APPLICATION NUMBER: JP 1999-300253

; PRIOR FILING DATE: 1999-08-27

; PRIOR APPLICATION NUMBER: JP 2000-118776

; PRIOR FILING DATE: 2000-01-11

; PRIOR APPLICATION NUMBER: JP 2000-183767

; PRIOR FILING DATE: 2000-05-02

; PRIOR APPLICATION NUMBER: JP 2000-241899

; PRIOR FILING DATE: 2000-06-09

; PRIOR APPLICATION NUMBER: 60/159, 590

; PRIOR FILING DATE: 1999-10-18

; PRIOR APPLICATION NUMBER: 60/183, 322

; PRIOR FILING DATE: 2000-02-17

; NUMBER OF SEQ ID NOS: 19025

; SOFTWARE: PatencIn Ver. 2.1

; SEQ ID NO 13028

; LENGTH: 353

; TYPE: PRT

; ORGANISM: Homo sapiens

US-09-629-469A-13028

```

Query Match          99.8%; Score 1868; DB 5; Length 353;
Best Local Similarity 99.7%; Pred. No. 6.9e-167;
Matches 352; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 MRYLPSVVLGTAPTYVLAMGWRLLSAFLPARFYQALDRLRYCYQSMVLEFFENYTG 60
    |||||||
DB 1 MRYLPSVVLGTAPTYVLAMGWRLLSAFLPARFYQALDRLRYCYQSMVLEFFENYTG 60
    |||||||
QY 61 VOILLYGLDLPKKNENITYLANHOSYVDWIVADILAIROMALGHVRVYVLEKGLKMLPLYGC 120
    |||||||
DB 61 VOILLYGLDLPKKNENITYLANHOSYVDWIVADILAIROMALGHVRVYVLEKGLKMLPLYGC 120
    |||||||
QY 121 YFAOHGCIYVKSASAKFNEKEMRNKLOSYYVDAGTAPMTLVIFPEGTRYNPQOTVLSASQAF 180
    |||||||
DB 121 YFAOHGCIYVKSASAKFNEKEMRNKLOSYYVDAGTAPMTLVIFPEGTRYNPQOTVLSASQAF 180
    |||||||
QY 181 AAORGLAVLKHVLTPIRIKATHVAFDCMKNYLDAIYDVTVVYEGKDGGRRESPTMTFEL 240
    |||||||
DB 181 AAORGLAVLKHVLTPIRIKATHVAFDCMKNYLDAIYDVTVVYEGKDGGRRESPTMTFEL 240
    |||||||
QY 241 CKECPKIHIDRIDKDVPEOEHRMRWLHERFEIKDKMLIEFYSPDERKKRPPGKS 300
    |||||||
```



Db 241 CKECPRIHHIDRIDKDVPEEOEHMRRLHEFEIKDKMLIEFYESPDPERRRRPPGKS 300  
QY 301 VNSKLSIKTLPMSLILSGLTAGMLMTDAGRKLIVNTWITGLGLMTVITKA 353  
Db 301 VNSKLSIKTLPMSLILSGLTAGMLMTDAGRKLIVNTWITGLGLMTVITKA 353

RESULT 3  
US-10-074-045-47  
; Sequence 47, Application US/10074045  
; GENERAL INFORMATION:  
; APPLICANT: Rosen et al.  
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies  
; FILE REFERENCE: PT221C1  
; CURRENT APPLICATION NUMBER: US/10/074,045  
; CURRENT FILING DATE: 2002-02-14  
; Prior Application removed - See File Wrapper or Palm  
; NUMBER OF SEQ ID NOS: 74  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 47  
; LENGTH: 269  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-074-045-47

Query Match 76.3%; Score 1427; DB 6; Length 269;  
Best Local Similarity 100.0%; Pred. No. 1,2e-125;  
Matches 269; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 85 TVDMIVADILAIIRONALGHVRYLKEGLKWLPLYGCFYFAOHGIIYKRSKAFNEKEKRNK 144  
Db 1 TVDMIVADILAIIRONALGHVRYLKEGLKWLPLYGCFYFAOHGIIYKRSKAFNEKEKRNK 60  
QY 145 LOSYVADGTPMYLVIPEEGTRYNPEQTKVLSASQAFPAOGLAVLKHVLTPTIKATHVAF 204  
Db 61 LOSYVADGTPMYLVIPEEGTRYNPEQTKVLSASQAFPAOGLAVLKHVLTPTIKATHVAF 120  
QY 205 DCKKNLDAIDVTVYVEGKDGQRRESPTMTTEFLCKECPKIHIDRIDKDVPEEOE 264  
Db 121 DCKKNLDAIDVTVYVEGKDGQRRESPTMTTEFLCKECPKIHIDRIDKDVPEEOE 180  
QY 265 HMRWLHEFEIKDKMLIEFYESPDPERRRRPPGKS VNSKLSIKTLPMSLILSGLTAGM 324  
Db 181 HMRWLHEFEIKDKMLIEFYESPDPERRRRPPGKS VNSKLSIKTLPMSLILSGLTAGM 240  
QY 325 LMTDAGRKLIVNTWITGLGLMTVITKA 353  
Db 241 LMTDAGRKLIVNTWITGLGLMTVITKA 269

RESULT 4  
US-09-935-625-11061  
; Sequence 11061, Application US/09935625  
; GENERAL INFORMATION:  
; APPLICANT: N. ALEXANDROV et al.  
; TITLE OF INVENTION: POLYNUCLEOTIDES, POLYPEPTIDES, CELLS, AND METHODS THEREOF CAPABLE  
; FILE REFERENCE: 2750-1481P  
; CURRENT APPLICATION NUMBER: US/09/935,625  
; CURRENT FILING DATE: 2001-08-24  
; NUMBER OF SEQ ID NOS: 33136  
; SEQ ID NO 11061  
; LENGTH: 375  
; TYPE: PRT  
; ORGANISM: Arabidopsis thaliana  
; FEATURE:  
; NAME/KEY: peptide  
; LOCATION: 1..375  
; OTHER INFORMATION: Ceres Seq. ID no. 1394370  
US-09-935-625-11061

Query Match 20.9%; Score 390.5; DB 5; Length 375;  
Best Local Similarity 26.9%; Pred. No. 3.9e-28;  
Matches 101; Conservative 78; Mismatches 128; Indels 69; Gaps 9;

QY 1 MRYLPSVVLGTAFTYVLAWG-----VWRLSAFLPARFYQLDRLCYVYOSMVLFFP 55  
Db 18 LRGIICLMLVSTAFPMILIFMGFLSAVLRLEFS-----IRYSRKCVSFFFGSWLALMPFLP 73  
QY 56 ENYTGVOILLGDLPRKNKEIITLANHOSVDMIVADILAIIRONALGHVRYLKEGLKWL 115  
Db 74 EKINKTKVIFSGDKVPCEDVLLIANHREVDMMYFMDLALRKGQGINIKYVLKSSLMKL 133  
QY 116 PLYGCFYFAOHGIIYKRSKAFNEKEKRNKLOSVDAGTPMYLVIPEEGTRYNPEQTKVLS 175  
Db 134 PLFGMAFHLEFIPVERRWVDEANLRQIVSSFKDRDRLMALLPFGDTYTEAKQ--- 190  
QY 176 ASQAFPAOGLAVLKHVLTPTIKATHVAFDCMKNYDAIDVTVYVEGKDGQRRESPT 235  
Db 191 RSKKFAENGLPIINNVLTPTKGFVSCLOELCSIDAVYDVITGYKTR----- 239  
QY 236 MTEFLCKECP-----KIHIDRIDKDVPEEOEHMRRLHEFEIKDKMLIE 283  
Db 240 -----CPSELDNYGIEPSEVHIIIRINLTQIPNDEKDIINAMLTFFQDKDLND 291  
QY 284 FYESPDERRRRPPGKS VNSKLSIKTLPMSLILSGLT-----AGMLMTDAGRKL 334  
Db 292 FYSN-----GHFPNDETEKEFNTKYLINCLAVIAFTTICTLTPSSMTF-----RIT 341  
QY 335 VNTWITGLGLMTV 350  
Db 342 VS-----LACYVLT 350

RESULT 5  
US-09-935-625-14197  
; Sequence 14197, Application US/09935625  
; GENERAL INFORMATION:  
; APPLICANT: N. ALEXANDROV et al.  
; TITLE OF INVENTION: POLYNUCLEOTIDES, POLYPEPTIDES, CELLS, AND METHODS THEREOF CAPA  
; FILE REFERENCE: 2750-1481P  
; CURRENT APPLICATION NUMBER: US/09/935,625  
; CURRENT FILING DATE: 2001-08-24  
; NUMBER OF SEQ ID NOS: 33136  
; SEQ ID NO 14197  
; LENGTH: 375  
; TYPE: PRT  
; ORGANISM: Arabidopsis thaliana  
; FEATURE:  
; NAME/KEY: peptide  
; LOCATION: 1..375  
; OTHER INFORMATION: Ceres Seq. ID no. 3441314  
US-09-935-625-14197

Query Match 20.9%; Score 390.5; DB 5; Length 375;  
Best Local Similarity 26.9%; Pred. No. 3.9e-28;  
Matches 101; Conservative 78; Mismatches 128; Indels 69; Gaps 9;

QY 1 MRYLPSVVLGTAFTYVLAWG-----VWRLSAFLPARFYQLDRLCYVYOSMVLFFP 55  
Db 18 LRGIICLMLVSTAFPMILIFMGFLSAVLRLEFS-----IRYSRKCVSFFFGSWLALMPFLP 73  
QY 56 ENYTGVOILLGDLPRKNKEIITLANHOSVDMIVADILAIIRONALGHVRYLKEGLKWL 115  
Db 74 EKINKTKVIFSGDKVPCEDVLLIANHREVDMMYFMDLALRKGQGINIKYVLKSSLMKL 133  
QY 116 PLYGCFYFAOHGIIYKRSKAFNEKEKRNKLOSVDAGTPMYLVIPEEGTRYNPEQTKVLS 175  
Db 134 PLFGMAFHLEFIPVERRWVDEANLRQIVSSFKDRDRLMALLPFGDTYTEAKQ--- 190  
QY 176 ASQAFPAOGLAVLKHVLTPTIKATHVAFDCMKNYDAIDVTVYVEGKDGQRRESPT 235  
Db 191 RSKKFAENGLPIINNVLTPTKGFVSCLOELCSIDAVYDVITGYKTR----- 239

Db 191 RSKFFAENGILPIINLVLPRTKGFVSCLOELSCSLDAVYDVITGYTR----- 239  
QY 236 MTEFLCKECP-----KIHIDRIDKKDVEEODHMRWLHERFEIKDKMLIE 283  
Db 240 -----CPSFLDNVYGIEPSEVHIHIRNLQIINQOEDIANMLNFTOLKDLND 291  
QY 284 FYSPDPERRRKRPFGKSVNSKLSIKTLPMSLLISGLT-----AGMLMTDAGRKL 334  
Db 292 FYSN-----GHFNEGTEKEFEFNKRYLINCIAVIAFTTICTHILTFPSSMIWF---RIV 341  
QY 335 VNTWITGTLGCLMVT 350  
Db 342 VS-----LACVYLT 350

RESULT 6  
US-09-935-625-14201  
; Sequence 14201, Application US/09935625

; GENERAL INFORMATION:  
; APPLICANT: N. ALEXANDROV et al.  
; TITLE OF INVENTION: POLYNUCLEOTIDES, POLYPEPTIDES, CELLS, AND METHODS THEREOF CAPABLE  
; FILE REFERENCE: 2750-1481P  
; CURRENT APPLICATION NUMBER: US/09/935, 625  
; CURRENT FILING DATE: 2001-08-24  
; NUMBER OF SEQ ID NOS: 33136  
; SEQ ID NO 14201  
; LENGTH: 375  
; TYPE: PRP  
; ORGANISM: Arabidopsis thaliana  
; NAME/KEY: peptide  
; LOCATION: 1..375  
; OTHER INFORMATION: Ceres Seq. ID no. 3441318  
US-09-935-625-14201

Query Match 20.9%; Score 390.5; DB 5; Length 375;  
Best Local Similarity 26.9%; Pred. No. 3.9e-28;  
Matches 101; Conservative 78; Mismatches 128; Indels 69; Gaps 9;

QY 1 MRYLLPSVILGTAPTYVLAMG-----VWRLLSAFLPAREYQALDDRLXCVOGSMVLF 55  
Db 18 LRGITCLMVLVSTAFPMMLIFWGLSAVVLRLFS---IRYSRKCVSFFFSWMLAMPFLF 73  
QY 56 ENYTGVOILLYGDLPKKNENIYLANHSTVDWIVADILAIRONALGHVRYVLKEGKML 115  
Db 74 EKIKTKVIFSGDKVPEDRVLLIANHRTEDWMYFMDLALRKGQIGNIKYVLSLMLK 133  
QY 116 PLYGCFYAOHGGLYVKRSKAFNEKEMRNKLOSVDAGTPMYLVIFPFGTRYINPEQTKVLS 175  
Db 134 PLFGMAFHLEFPIVERRWEDEANLRQIVSSFQDPDALMLALFPEGDTYTEAKCO--- 190  
QY 176 ASQAFAROGIAVLKHYLTPRIKATHVAFDCMKNYLDAIDVYVYEGKDDGGRRRESPT 235  
Db 191 RSKFFAENGILPIINLVLPRTKGFVSCLOELSCSLDAVYDVITGYTR----- 239  
QY 236 MTEFLCKECP-----KIHIDRIDKKDVEEODHMRWLHERFEIKDKMLIE 283  
Db 240 -----CPSFLDNVYGIEPSEVHIHIRNLQIINQOEDIANMLNFTOLKDLND 291  
QY 284 FYSPDPERRRKRPFGKSVNSKLSIKTLPMSLLISGLT-----AGMLMTDAGRKL 334  
Db 292 FYSN-----GHFNEGTEKEFEFNKRYLINCIAVIAFTTICTHILTFPSSMIWF---RIV 341  
QY 335 VNTWITGTLGCLMVT 350  
Db 342 VS-----LACVYLT 350

RESULT 7  
US-09-935-625-14254  
; Sequence 14254, Application US/09935625

; GENERAL INFORMATION:  
; APPLICANT: N. ALEXANDROV et al.  
; TITLE OF INVENTION: POLYNUCLEOTIDES, POLYPEPTIDES, CELLS, AND METHODS THEREOF CAPA  
; FILE REFERENCE: 2750-1481P  
; CURRENT APPLICATION NUMBER: US/09/935, 625  
; CURRENT FILING DATE: 2001-08-24  
; NUMBER OF SEQ ID NOS: 33136  
; SEQ ID NO 14254  
; LENGTH: 375  
; TYPE: PRP  
; ORGANISM: Arabidopsis thaliana  
; NAME/KEY: peptide  
; LOCATION: 1..375  
; OTHER INFORMATION: Ceres Seq. ID no. 3443248  
US-09-935-625-14254

Query Match 20.9%; Score 390.5; DB 5; Length 375;  
Best Local Similarity 26.9%; Pred. No. 3.9e-28;  
Matches 101; Conservative 78; Mismatches 128; Indels 69; Gaps 9;

QY 1 MRYLLPSVILGTAPTYVLAMG-----VWRLLSAFLPAREYQALDDRLXCVOGSMVLF 55  
Db 18 LRGITCLMVLVSTAFPMMLIFWGLSAVVLRLFS---IRYSRKCVSFFFSWMLAMPFLF 73  
QY 56 ENYTGVOILLYGDLPKKNENIYLANHSTVDWIVADILAIRONALGHVRYVLKEGKML 115  
Db 74 EKIKTKVIFSGDKVPEDRVLLIANHRTEDWMYFMDLALRKGQIGNIKYVLSLMLK 133  
QY 116 PLYGCFYAOHGGLYVKRSKAFNEKEMRNKLOSVDAGTPMYLVIFPFGTRYINPEQTKVLS 175  
Db 134 PLFGMAFHLEFPIVERRWEDEANLRQIVSSFQDPDALMLALFPEGDTYTEAKCO--- 190  
QY 176 ASQAFAROGIAVLKHYLTPRIKATHVAFDCMKNYLDAIDVYVYEGKDDGGRRRESPT 235  
Db 191 RSKFFAENGILPIINLVLPRTKGFVSCLOELSCSLDAVYDVITGYTR----- 239  
QY 236 MTEFLCKECP-----KIHIDRIDKKDVEEODHMRWLHERFEIKDKMLIE 283  
Db 240 -----CPSFLDNVYGIEPSEVHIHIRNLQIINQOEDIANMLNFTOLKDLND 291  
QY 284 FYSPDPERRRKRPFGKSVNSKLSIKTLPMSLLISGLT-----AGMLMTDAGRKL 334  
Db 292 FYSN-----GHFNEGTEKEFEFNKRYLINCIAVIAFTTICTHILTFPSSMIWF---RIV 341  
QY 335 VNTWITGTLGCLMVT 350  
Db 342 VS-----LACVYLT 350

RESULT 8  
US-09-935-625-23009  
; Sequence 23009, Application US/09935625

; GENERAL INFORMATION:  
; APPLICANT: N. ALEXANDROV et al.  
; TITLE OF INVENTION: POLYNUCLEOTIDES, POLYPEPTIDES, CELLS, AND METHODS THEREOF CAPA  
; FILE REFERENCE: 2750-1481P  
; CURRENT APPLICATION NUMBER: US/09/935, 625  
; CURRENT FILING DATE: 2001-08-24  
; NUMBER OF SEQ ID NOS: 33136  
; SEQ ID NO 23009  
; LENGTH: 375  
; TYPE: PRP  
; ORGANISM: Arabidopsis thaliana  
; NAME/KEY: peptide  
; LOCATION: 1..375  
; OTHER INFORMATION: Ceres Seq. ID no. 3441318  
US-09-935-625-23009

Query Match	20.9%;	Score 390.5;	DB 5;	Length 375;
Best Local Similarity	26.9%;	Pred. No. 3.9e-28;		
Matches 101; Conservative	78;	Mismatches 128;	Indels 69;	Gaps 9;

```

QY      1 MRVLLSPVLLGTGPTVYLIAG-----VNRLLSAFLPARFYQALDDRLCYOSOMVLEFF 55
      1 : :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: ::
Db      18 LRGIICLAWLVSTFPMALIFMGFLSAVYLRLE-----IRSRKCVSEFFPGSMLWMPLEF 73
      18 : :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: ::
QY      56 ENYTGVOILLLEGDLPKKNENIYLANHOSQTPMDIADILAIRONALCHVRVLEKEGKWL 119
      56 : :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: ::
Db      74 EKIKTKVIFESGDVPCEDRRLLIANHRTVEDMWYTFMDLARKQIQIKYVLSKLML 133
      74 : :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: ::
QY      116 PLVGYCYFAHGQIYVKSASAKFNEKEMRNKIQSQSYDACTPMYLVTFPEGSTRYNEQOTVLS 175
      116 : :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: ::
Db      134 PLFMFAHFLFEFIEVERMEWEVDENLRQIVSSFKDPDRDALMLALFPEGTDTYEAKQ--- 190
      134 : :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: ::
QY      176 ASQAFNAQRGLAVLKHVLTPIKATIHAFDCMKRYLDALDYVYVYEGKDDGGGRRESPT 233
      176 : :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: ::
Db      191 RSKFPAENGGPIRLINVLPRTKGFVSCLOELSCSDAVVDYVTIGYKTR----- 238
      191 : :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: ::
QY      236 MTEFLCKECP-----KIHIDRLDKDVEEEOGHMRWVHERPEIKDKMLIE 288
      236 : :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: ::
Db      240 -----CPSEFLDNYGIERSEVHHIIRKRLNLTQIENQEKDIAMLMNTPQOLKDLND 291
      240 : :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: ::
QY      284 FYESPDPERRRRRPFQGSVNSKLSIKTKLPSMLISGLT-----AGMLINDAGRKLY 334
      284 : :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: ::
Db      292 FYSN-----GHFNEGEGKEFEKFNKRYLINCLAVIAFTVITCTHLEFPSSMIWF---RIY 341
      292 : :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: ::
QY      335 VNTWITGTLGCLMVT 350
      335 : :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: ::
Db      342 VS-----LACVYILT 350
      342 : :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: ::

```

RESULT 9  
US-09-935-625-23025  
; Sequence 23025, Application US/09935625

Query Match	20.9%;	Score 390.5;	DB 5;	Length 375;
Best Local Similarity	26.9%;	Pred. No. 3.9e-28;		
Matches 101;	Conservative 78;	Mismatches 128;	Indels 69;	Gaps 9;

Qy	1	MRYLPSVLLGTPPYVLAM-----VWRLSLFLPRLFYQALDRLXYCSQSMVLEFF	55
Qy	1	::: :::: :::: :::: :::: :::: :::: :::: :::: :::: :::: :::: :::: :::: ::::	
Db	18	LRGILITLAVLSTAPMLITFEGFLSAVLRLEF---IRSRKCVSEFFESQWALMPFLF	73
Qy	56	ENYGVQVOLLIGDLPKKNENIYLANHQSOTYVADILAIRONALHVRVLYKEGKWL	115
Qy	1	::: :::: :::: :::: :::: :::: :::: :::: :::: :::: :::: :::: :::: ::::	
Db	74	EIKIKTKYIEGDAVPCEDRVLILANHRTVEYDMYTFMDLARKQIGIKITVYKSSLMKL	133
Qy	116	PLVGYCPAOGHGIVYKSAKENEKEMRNKLOSVDAGTPLYLTFPEGSTRYNEQOTVLS	175
Qy	1	::: :::: :::: :::: :::: :::: :::: :::: :::: :::: :::: :::: :::: ::::	
Db	134	PLFGMAFLTFEIEYVERRWEVDANLRQIYSSFPDPRDALMLALFPEGSTDTYTAKKQ---	190
Qy	176	ASQAFAMQRGLAVLKHVLTLPRIKTHVAFDMCKMYLDAIYDVTVVYEGKDDGGGRRESPT	235

```

Db      191  RSKFPAENGLPIANNVLPRTKGFVSCLOECSLDAYDVITGYTR-----239
          | ||| : || : || || : ||| |||| : | :
QY      236  MTEFLCKECP-----KIHIDRIDKKDVEEQEHMRKHERFEIKDMLIE 283
          || : |||| || : : ||| : || : |||| : | :
Db      240  -----CPFELNDVYGIEPSEVHHIRNRNLQTITRQEKDINAMLMNTFQKDLND 291
          || : |||| || : : ||| : || : |||| : | :
QY      284  FYSPDEPRRRRPFQGSVNSKLSIKTLPMSLISGLT-----AGMLMTDAGRKL 334
          || : || : || : || : || : || : || : || : || :
Db      292  FYSN-----GHFNEGEEKEFNKKYLINCLAVIAFTTICTHLPFFSSIMF---RIY 341
          || : || : || : || : || : || : || : || : || :
QY      335  VNTWLYGTLIGLQAMVT 350
          | : |||| :
Db      342  VS-----LACVYLT 350
          || : |||| :

```

```

RESULT 10
US-09-935-625-11062
: Sequence 11062, Application US/09935625
: GENERAL INFORMATION:
: APPLICANT: N. ALEXANDROV et al.
: TITLE OF INVENTION: POLYNUCLEOTIDES, POLYPEPTIDES, CELLS, AND METHODS THEREOF CAPABLE OF
: FILE REFERENCE: 2750-1481P
: CURRENT APPLICATION NUMBER: US/09/935,625
: CURRENT FILING DATE: 2001-08-24
: NUMBER OF SEQ ID NOS: 33136
: SEQ ID NO 11062
: LENGTH: 351
: TYPE: PRT
: ORGANISM: Arabidopsis thaliana
: FEATURE:
: NAME/KEY: peptide
: LOCATION: 1..351
: OTHER INFORMATION: Ceres Seq. ID no. 1394371
: US-09-935-625-11062

```

Query Match	20.7%	Score 387.5;	DB 5;	Length 351;
Best Local Similarity	27.1%;	Pred. No. 6.7e-28;		
Matches 100;	Conservative 75;	Mismatches 125;	Indels 69;	Gaps 9;

```

QY 8 VLLGTAPTYLANG-----VMRLSLAPLAREFOALDRLRYCYOSMVFEEFFENYGVQ 62
Db 1 MVLVSTAFMMLHMGFSLAVVLLFS-----IRSRKCVSFFEGSMALMPLPELEKINKTK 56
QY 63 ILLYGDPKKNKENITYLANHOSYDWMIVADILAIRQNALGHVRVLEEGSKWLPYGCYF 1222
   :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: ::
Db 57 VIFSGDKPCGCDRVLILIANHTEVDWMYFMDLALRKQOINIKRYVLSSLMKLPLFGWAF 116
QY 123 AQHGGIYVKSAKTNEKEMRNKLOSVDAGTPLYIIFPEGSTRYNPQUTRYLASAQAF 1822
   :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: ::
Db 117 HLFEPIVEREMVEDENLRQIYSSFPDPRDALMALFPEGTGYTEAKC-----RSKKFAA 173
QY 183 QRLGLAVLKHVLTPLRIKATNHAFOCDMKYVLAIVYTVYVYEGSKDGGRRSPMTTERLCK 242
   :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: ::
Db 174 ENGLPLINLVLLPRTKGFVSCLOELSSCLDAVDVYVTLGYIKTR----- 215
QY 243 ECP-----KIHIIIDRIDDKDVPDEEGEHMRMRWHERERIKDKMLIEYESPD 2908
   :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: ::
Db 216 -CPSELDNVYIGIEPSEVHIHRIIRNLQIPIQOENDINAMLMNFQLOLDOLLNDFYSN- 271
QY 291 ERRRRPPEKSVNSKLSIKTKPLSMILISGLT-----AGMLMTDAGRKLYVNIWYIG 3411
   :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: ::
Db 272 ---GHFPEEGTEKEBFNKKYVLINCLAVIAFTTICTHLEFFSSMIWF-----RIYVS- 319
QY 342 TLLGCIWYT 350
Db 320 -LACVLYLT 326

```

RESULT 11  
US-09-935-625-14198

```
; Sequence 14198, Application US/09935625
; GENERAL INFORMATION:
; APPLICANT: N. ALEXANDROV et al.
; TITLE OF INVENTION: POLYNUCLEOTIDES, POLYPEPTIDES, CELLS, AND METHODS THEREOF CAPABLE
; OF INVENTION: MODULATING VARIOUS RESPONSES
; FILE REFERENCE: 2750-1481P
; CURRENT APPLICATION NUMBER: US/09/935, 625
; CURRENT FILING DATE: 2001-08-24
; NUMBER OF SEQ ID NOS: 33136
; SEQ ID NO 14198
; LENGTH: 351
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; NAME/KEY: peptide
; LOCATION: 1..351
; OTHER INFORMATION: Ceres Seq. ID no. 3441315
US-09-935-625-14198
```

```
Query Match 20.7%; Score 387.5; DB 5; Length 351;
Best Local Similarity 27.1%; Pred. No. 6.7e-28;
Matches 100; Conservative 75; Mismatches 125; Indels 69; Gaps 9;
```

```
QY 8 VLLGTAPTVYLVWG-----VWRLSAFLPARFYQALDRLCVYOSMVLFFPENTGVQ 62
   :||: || :||: || :||: || :||: || :||: || :||: || :||: || :||: ||
Db 1 MVLVSTAFMMLIFWGLSAVVLRLFS---IRSRKCVSFFFGSWLALWPFLEKINKTK 56
   :||: || :||: || :||: || :||: || :||: || :||: || :||: || :||: ||
QY 63 ILLYGLPKKKNIIYLANHSTVDIVADILAIRONALGHVRYVLEKGLMPLTGYCF 122
   :||: || :||: || :||: || :||: || :||: || :||: || :||: || :||: ||
Db 57 VIFSGDKVPCEDRVLLIANHRTVDMMYFWDLARKGQIGNITYVLSKSLMKPLPLGMAF 116
   :||: || :||: || :||: || :||: || :||: || :||: || :||: || :||: ||
QY 123 AOHGIIYKRSKAFNEKEMRNKLOSVDAGTPMYLVIFPEGRYRNPEQTKVLSASQAF 182
   :||: || :||: || :||: || :||: || :||: || :||: || :||: || :||: ||
Db 117 HLEFETIVERRWEVDANLQIVSFKPRDALMLALFPEGDTYTEAKQ---RSKKFAA 173
   :||: || :||: || :||: || :||: || :||: || :||: || :||: || :||: ||
QY 183 QRLAVLKHVLTPIKATHTVAFCMKNYLDATIDVTVYVEGKDDGGQRESPTMTFELCK 242
   :||: || :||: || :||: || :||: || :||: || :||: || :||: || :||: ||
Db 174 ENGLPILNNVLRPTKGFVSCLOELSCSDAVYDVITGYKTR-----RIVVS----- 215
   :||: || :||: || :||: || :||: || :||: || :||: || :||: || :||: ||
QY 243 ECP-----KIHIIHIDRIDKKDVEDEGHRMRMLHERFETIKDKMLIEFESPPD 290
   :||: || :||: || :||: || :||: || :||: || :||: || :||: || :||: ||
Db 216 -CPSFLDNYGIPSEVHIIRINLTQIPNOEKDINAMLMNTFQLDKDLNDFYSN--- 271
   :||: || :||: || :||: || :||: || :||: || :||: || :||: || :||: ||
QY 291 ERRKRPFGKSVNSKLSIKKTLPSMLISGLT-----AGMLMTDAGRKLVTNTWIYG 341
   :||: || :||: || :||: || :||: || :||: || :||: || :||: || :||: ||
Db 272 ---GHPNECTEKEFMTKKYLINCLAVIAFTTICTHLPFSSMIWF---RIVVS----- 319
   :||: || :||: || :||: || :||: || :||: || :||: || :||: || :||: ||
QY 342 TLIGCLMWT 350
   :||: || :||: || :||: || :||: || :||: || :||: || :||: || :||: ||
Db 320 --LACVYLT 326
   :||: || :||: || :||: || :||: || :||: || :||: || :||: || :||: ||
```

```
RESULT 12
US-09-935-625-14202
; Sequence 14202, Application US/09935625
; GENERAL INFORMATION:
; APPLICANT: N. ALEXANDROV et al.
; TITLE OF INVENTION: POLYNUCLEOTIDES, POLYPEPTIDES, CELLS, AND METHODS THEREOF CAPABLE
; OF INVENTION: MODULATING VARIOUS RESPONSES
; FILE REFERENCE: 2750-1481P
; CURRENT APPLICATION NUMBER: US/09/935, 625
; CURRENT FILING DATE: 2001-08-24
; NUMBER OF SEQ ID NOS: 33136
; SEQ ID NO 14202
; LENGTH: 351
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; NAME/KEY: peptide
; LOCATION: 1..351
; OTHER INFORMATION: Ceres Seq. ID no. 3441319
US-09-935-625-14202
```

```
Query Match 20.7%; Score 387.5; DB 5; Length 351;
Best Local Similarity 27.1%; Pred. No. 6.7e-28;
Matches 100; Conservative 75; Mismatches 125; Indels 69; Gaps 9;
```

```
QY 8 VLLGTAPTVYLVWG-----VWRLSAFLPARFYQALDRLCVYOSMVLFFPENTGVQ 62
   :||: || :||: || :||: || :||: || :||: || :||: || :||: || :||: ||
Db 1 MVLVSTAFMMLIFWGLSAVVLRLFS---IRSRKCVSFFFGSWLALWPFLEKINKTK 56
   :||: || :||: || :||: || :||: || :||: || :||: || :||: || :||: ||
QY 63 ILLYGLPKKKNIIYLANHSTVDIVADILAIRONALGHVRYVLEKGLMPLTGYCF 122
   :||: || :||: || :||: || :||: || :||: || :||: || :||: || :||: ||
Db 57 VIFSGDKVPCEDRVLLIANHRTVDMMYFWDLARKGQIGNITYVLSKSLMKPLPLGMAF 116
   :||: || :||: || :||: || :||: || :||: || :||: || :||: || :||: ||
QY 123 AOHGIIYKRSKAFNEKEMRNKLOSVDAGTPMYLVIFPEGRYRNPEQTKVLSASQAF 182
   :||: || :||: || :||: || :||: || :||: || :||: || :||: || :||: ||
Db 117 HLEFETIVERRWEVDANLQIVSFKPRDALMLALFPEGDTYTEAKQ---RSKKFAA 173
   :||: || :||: || :||: || :||: || :||: || :||: || :||: || :||: ||
QY 183 QRLAVLKHVLTPIKATHTVAFCMKNYLDATIDVTVYVEGKDDGGQRESPTMTFELCK 242
   :||: || :||: || :||: || :||: || :||: || :||: || :||: || :||: ||
Db 174 ENGLPILNNVLRPTKGFVSCLOELSCSDAVYDVITGYKTR-----RIVVS----- 215
   :||: || :||: || :||: || :||: || :||: || :||: || :||: || :||: ||
QY 243 ECP-----KIHIIHIDRIDKKDVEDEGHRMRMLHERFETIKDKMLIEFESPPD 290
   :||: || :||: || :||: || :||: || :||: || :||: || :||: || :||: ||
Db 216 -CPSFLDNYGIPSEVHIIRINLTQIPNOEKDINAMLMNTFQLDKDLNDFYSN--- 271
   :||: || :||: || :||: || :||: || :||: || :||: || :||: || :||: ||
QY 291 ERRKRPFGKSVNSKLSIKKTLPSMLISGLT-----AGMLMTDAGRKLVTNTWIYG 341
   :||: || :||: || :||: || :||: || :||: || :||: || :||: || :||: ||
Db 272 ---GHPNECTEKEFMTKKYLINCLAVIAFTTICTHLPFSSMIWF---RIVVS----- 319
   :||: || :||: || :||: || :||: || :||: || :||: || :||: || :||: ||
QY 342 TLIGCLMWT 350
   :||: || :||: || :||: || :||: || :||: || :||: || :||: || :||: ||
Db 320 --LACVYLT 326
   :||: || :||: || :||: || :||: || :||: || :||: || :||: || :||: ||
```

```
RESULT 13
US-09-935-625-14255
; Sequence 14255, Application US/09935625
; GENERAL INFORMATION:
; APPLICANT: N. ALEXANDROV et al.
; TITLE OF INVENTION: POLYNUCLEOTIDES, POLYPEPTIDES, CELLS, AND METHODS THEREOF CAPA
; OF INVENTION: MODULATING VARIOUS RESPONSES
; FILE REFERENCE: 2750-1481P
; CURRENT APPLICATION NUMBER: US/09/935, 625
; CURRENT FILING DATE: 2001-08-24
; NUMBER OF SEQ ID NOS: 33136
; SEQ ID NO 14255
; LENGTH: 351
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; NAME/KEY: peptide
; LOCATION: 1..351
; OTHER INFORMATION: Ceres Seq. ID no. 3443249
US-09-935-625-14255
```

```
Query Match 20.7%; Score 387.5; DB 5; Length 351;
Best Local Similarity 27.1%; Pred. No. 6.7e-28;
Matches 100; Conservative 75; Mismatches 125; Indels 69; Gaps 9;
```

```
QY 8 VLLGTAPTVYLVWG-----VWRLSAFLPARFYQALDRLCVYOSMVLFFPENTGVQ 62
   :||: || :||: || :||: || :||: || :||: || :||: || :||: || :||: ||
Db 1 MVLVSTAFMMLIFWGLSAVVLRLFS---IRSRKCVSFFFGSWLALWPFLEKINKTK 56
   :||: || :||: || :||: || :||: || :||: || :||: || :||: || :||: ||
QY 63 ILLYGLPKKKNIIYLANHSTVDIVADILAIRONALGHVRYVLEKGLMPLTGYCF 122
   :||: || :||: || :||: || :||: || :||: || :||: || :||: || :||: ||
Db 57 VIFSGDKVPCEDRVLLIANHRTVDMMYFWDLARKGQIGNITYVLSKSLMKPLPLGMAF 116
   :||: || :||: || :||: || :||: || :||: || :||: || :||: || :||: ||
QY 123 AOHGIIYKRSKAFNEKEMRNKLOSVDAGTPMYLVIFPEGRYRNPEQTKVLSASQAF 182
   :||: || :||: || :||: || :||: || :||: || :||: || :||: || :||: ||
Db 117 HLEFETIVERRWEVDANLQIVSFKPRDALMLALFPEGDTYTEAKQ---RSKKFAA 173
   :||: || :||: || :||: || :||: || :||: || :||: || :||: || :||: ||
```

[illegible]

RESULT 14  
US-09-935-625-23010  
; Sequence 23010, Application US/09935625  
Copyright © 2010, Application US/09935625

```

? APPLICANT: N. ALEXANDROV et al.
? TITLE OF INVENTION: POLYNUCLEOTIDES, PEPTIDES, CELLS, AND METHODS THEREOF CAPABLE
? TITLE OF INVENTION: MODULATING VARIOUS RESPONSES
? FILE REFERENCE: 2/50-1481P
? CURRENT APPLICATION NUMBER: US/09/935,625
? CURRENT FILING DATE: 2001-08-24
? NUMBER OF SEQ ID NOS: 33136
? SEQ ID NO 23010
? LENGTH: 351
? TYPE: PRT
? ORGANISM: Arabidopsis thaliana
? FEATURE:
? NAME/KEY: peptide
? LOCATION: 1..351
? OTHER INFORMATION: Ceres Seq. ID no. 3441319
? US-09-935-625-23010

```

```

US-09-935-625 23026
; Sequence 23026, Application US/09935625
; GENERAL INFORMATION:
; APPLICANT: N. ALEXANDROV et al.
; TITLE OF INVENTION: POLYNUCLEOTIDES, POLYPEPTIDES, CELLS, AND METHODS THEREOF CAPAB
; TITLE OF INVENTION: MODULATING VARIOUS RESPONSES
; FILE REFERENCE: 2750-1481P
; CURRENT APPLICATION NUMBER: US/09/935,625
; CURRENT FILING DATE: 2001-08-24
; NUMBER OF SEQ ID NOS: 33136
; SEQ ID NO 23026
; LENGTH: 351
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; NAME/KEY: peptide
; LOCATION: 1..351
; OTHER INFORMATION: Ceres Seq. ID no. 3443249
US-09-935-625-23026

```

	Query Match	20.7%	Score 387.5;	DB 5;	Length 351;	
	Best Local Similarity	27.1%	Pred. No. 6,7e-28;			
	Matches 100;	Conservative 75;	Mismatches 125;	Indels 69;	Gaps 9;	
Oy	8 VVLGTAPTYVLAWG----	VWRLLSAFLPARFYALDRLRYCYQSYMLFFEFENYTGQV	62			
	:::::	:::::	:::::			
Db	1 MVLSTAFMLLIFMFGLSAAIVLRFES---	IIRSKVCSEFFGSGWLALPFLPEFKINRKT	56			
	:::::	:::::	:::::			
Oy	63 ILLVDGLPKRNKNIILYNHOSHYVMIAVDIALIROMNLGHVRVYLKBELMKLPYGCFF	122				
	:::::	:::::	:::::			
Db	57 VFESGDGVPCEDERVULLIAHRTVEVMVFWDALRKGGQGNIKYLVKSLSKLPLFEGMAF	116				
	:::::	:::::	:::::			
Oy	123 AQHGSIYYRSAKAFPEKEEMRNKLOSIVADAGTMYLVIFEEPTGRNYPEQTGYVSASQAFA	182				
	:::::	:::::	:::::			
Db	117 HLFEFIIPERMREVDDEANLRQLTVSSPKDRDALMWLALPEEGTDYTEAKQQ--RSKKFAA	173				
	:::::	:::::	:::::			
Oy	183 QRGSLAVLKHVLTPRIRKAITHVAFCDKKNYIDALDYVTYVTEGKDQSGQRRESPTMTIEFLCK	242				
	:::::	:::::	:::::			
Db	174 ENGSLFINNVLLPRTKGFVSCLOEJLSCSDIIAYDVDTIGYKTR-----	215				
	:::::	:::::	:::::			
Oy	243 ECP-----KIHHIDRIDKODVPDEOEHRMRRLHREFIKDKMLIEFPSPDP	290				
	:::::	:::::	:::::			
Db	216 -CPSFLDNVYGIEBSEVNHRIIRINULTQIPNOEKDIMAMLMNTFOQLKOOLDLDFPSN---	271				
	:::::	:::::	:::::			
Oy	291 ERRKRPFCKSVNSKLSIKKTLTPSMILISGLT-----ACGLMTDAGRKLYVNWTWIYG	341				
	:::::	:::::	:::::			
Db	272 ---GHFPNGTEKEKENTKYLLINCLAVIAFTTICTHLHFESSMIWF---RIYVS-----	319				
	:::::	:::::	:::::			
Oy	342 TLLGGLMTY 350					
	:::::	:::::	:::::			
Db	320 --LACVIYLT 326					

Search completed: August 28, 2002, 11:15:41  
Job time: 555 sec



GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 28, 2002, 11:04:46 ; Search time 39.13 Seconds  
(without alignments)  
866.842 Million cell updates/sec

Title: US-09-853-526-4  
Perfect score: 1871

Sequence: 1 MRYLPSVVLGTAPTYVLA.....YVNWIVYGLLGLMTVTKA 353

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283138 segs, 96089334 residues

Total number of hits satisfying chosen parameters: 283138

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : PIR71:\*  
1: pir1:\*  
2: pir2:\*  
3: pir3:\*  
4: pir4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	453	24.2	918	2	T34057
2	398.5	21.3	393	2	B96780
3	328	17.5	374	2	S52645
4	307	16.4	377	2	S60478
5	263	14.1	376	2	D96550
6	249.5	13.3	350	2	T40466
7	242.5	13.0	344	2	T31913
8	239	12.8	311	2	T07936
9	238.5	12.7	310	2	T06755
10	231.5	12.4	397	2	S45900
11	221.5	11.8	396	2	S54641
12	213.5	11.4	523	2	T25998
13	212	11.3	439	2	T23689
14	182	9.7	391	2	T15366
15	178	9.5	295	2	B83587
16	154	8.2	310	2	S40808
17	154	8.2	310	2	A91227
18	154	8.2	310	2	G66073
19	152.5	8.2	304	2	B83541
20	139.5	7.5	285	2	T15252
21	135.5	7.2	302	2	A10950
22	131.5	7.0	363	2	T20608
23	130	6.9	294	2	A48600
24	120.5	6.4	303	2	D82371
25	118	6.3	281	2	S60477
26	114.5	6.1	247	2	G72223
27	114	6.1	266	2	G73863
28	112	6.0	258	2	G83101
29	111.5	6.0	488	1	OCBEHS

30	111	5.9	377	1	A44216	major envelope ant
31	109.5	5.9	488	2	T44030	alkaline exonuclease
32	109.5	5.9	488	2	T44215	alkaline exonuclease
33	109	5.8	240	2	D64089	1-acetylglucosyl-3-p
34	108	5.8	261	2	C82067	probable 1-acetylglucosyl
35	104	5.6	240	2	D64688	probable 1-acetylglucosyl
36	104	5.6	267	2	AC3499	1-acetyl-sn-glycerol
37	103	5.5	243	2	AC0083	1-acetylglucosyl-3-p
38	103	5.5	266	2	B98283	hypothetical prote
39	103	5.5	266	2	AC3000	1-acetyl-sn-glycerol
40	102	5.5	212	2	G85357	hypothetical prote
41	101	5.4	262	2	T22599	1-acetylglucosyl-3-p
42	101	5.4	323	2	C83940	sugar ABC transpor
43	99.5	5.3	237	2	B71827	probable 1-acetylglucosyl
44	99	5.3	225	2	B45582	probable 1-acetylglucosyl
45	99	5.3	243	2	B71706	probable 1-acetylglucosyl

## ALIGNMENTS

RESULT	1	Score	DB 2	Length	918
T34057	hypothetical protein F28B3.5 - <i>Caenorhabditis elegans</i>	32.9%	Pred. No. 9.3e-31;		
C:Species:	<i>Caenorhabditis elegans</i>				
C:Date:	29-Oct-1999	#sequence_revision	29-Oct-1999	#text_change	29-Oct-1999
C:Accession:	T34057				
R:Geisel, C.; Kramer, J.; Smith, A.					
A:Submitted to the EMBL Data Library, May 1997					
A:Description:	The sequence of <i>C. elegans</i> cosmid F28B3.				
A:Reference number:	221469				
A:Accession:	T34057				
A>Status:	preliminary; translated from GB/EMBL/DBJ				
A:Molecule type:	DNA				
A:Residues:	1-918 <GEI>				
A:Cross-References:	EMBL:AF003136; PIDN:AA93636.1; GSPDB:GN00019; CESP:F28B3.5				
A:Experimental source:	strain Bristol N2; clone F28B3				
C:Genetics:					
A:Gene:	CESP:F28B3.5				
A:Map position:	1				
A:Introns:	85/3; 129/3; 235/3; 418/2; 482/3; 532/3; 736/3; 829/3; 850/2				
Query Match	24.2%	Score	453	DB 2	Length 918;
Best Local Similarity	32.9%	Pred. No.	9.3e-31;		
Matches	127; Conservative	72; Mismatches	139; Indels	48; Gaps	14;
QY	1	MRYLPSVVLGTAPTYVLA	MGVRLSAFLPAREQALDRLCVYSMLVFPE	56	
DB	544	LRRIICSLSLSMVFPFACALIVGV	SWIVRHAQQQLNMNLYKSYMLCLFVPE	599	
QY	57	NTGVOILLYGDLK--NK----	ENIIYLANHOSIVDMIVADILAIRONALGH--	VRRV	107
DB	600	NSGVEIYHGTNEEVNKTGKRENVAMISNOSNDWIIPVLAARHSDQGEQAFRYM		659	
QY	108	IKREGKLMLPGCYFQHOHGIVYKRSKFNEMRKKSQSYDAGPMTLVIFPESTRYN	167		
DB	660	VKNSHVLVPMFGWYIFOHGIVYRRGEFTGAPVLRLQLMLNSDPYMLLIFPESTRNS	719		
QY	168	PEOTKVLASQAFAAQRGLAVLKHVLT	PRIKATVAFDCMKNYLDAIVDVTVVY--	EGKDD	226
DB	720	AKKHLLSSNRLEKSGNQPMQNVLC	PRSGGQLALDMLNST--LDAIVDVTVVYGGMRD	778	
QY	227	GC-----	QRRESPTMTEFLC--	KECPKIHIDRIDKKDVPEDQENR	267
DB	779	LGINLTIHYNLIIITYRMAERGLAP	GMFCGSOQFOKLHILHIDRPIDEVPKALELR	838	
QY	268	RMHHEFEIKDKMLIEFYSPPDERKRR	PKGSVNSKLSIKKTLPEMLLS--	GLTAGMLM	326
DB	839	TWTEFEFTKERTIDEFY--SEKSTGSALP	-----	CVPLSQTLPTSLFFSALALAPFS	891
QY	327	TDAGRLVYVNWIVYGLLGLMTVTKA		352	

Db 892 RTIGR-IYLLTIASSPLL-TAMLHIR 915

## RESULT 2

hypothetical protein F9E10.13 [imported] - Arabidopsis thaliana

C:Species: Arabidopsis thaliana (mouse-ear cress)

C:Date: 02-Mar-2001 #sequence\_revision 02-Mar-2001 #text\_change 31-Mar-2001

C:Accession: B96780

R:Theologas, A.; Ecker, J.R.; Palm, C.J.; Pederspiel, N.A.; Kaul, S.; White, O.; Alonso, Chn, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.; aneun, N.F.; Hughes, B.; Hutzar, L.

Nature 408, 816-820, 2000

A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.C.A.; Li, J.H.; Li, Y.; Liu, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maitl, R.; Marziani, Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.

A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon, ker, M.; Wu, D.; Xu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.

A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.

A:Reference number: A86141; MUID:21016719

A:Accession: B96780

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1393 <STO>

A:Cross-references: GB:AE005173; NID:g6646762; PIDN:AAF21074.1; GSPDB:GN00141

C:Genetics:

A:Gene: F9E10.13

A:Map position: 1

Query Match 21.3%; Score 398.5; DB 2; Length 393;  
Best Local Similarity 29.8%; Pred. No. 1.5e-26;  
Matches 97; Conservative 68; Mismatches 117; Indels 43; Gaps 7;

1 MRYLPSVLLGTAAPYVLAAGVRLASFLPARFYQALDRLCYVQSMV----- 52

23 LRLGLMLVLSLAFMLL-----YFAP---IALGLRLSVQSKVSLIGLW 70

53 -----FEFENVYGVQILLYDLPKKNKIITILANHOSTVDMIVADILAIRONALGHVRY 107

71 LALMPYLFETVNGTGVVFSDDIIPVEKRVLLINHRTREVDMYLMIALKKGGLYKYV 130

108 LKGLKMLPRIGCYFAOHGIIYKRSKFNKEMKRNKLOSVDAGTPMYLVIPPEGRTYN 167

131 LKSLKMLPFGMGFHLERIPVEKRVDEPVLQMLSSFKDQPEPLMLALPEEGDFT 190

168 PEQTKVLSQAFAAORGAVLKLHVLPRIKATHVAFDCMKNYLDATYDVYVEGKDDG 227

191 EEKCK---RSQKFAEYGLPALSNVLPKTRGVCLEVLHNSLDAYDITIAKPR--- 244

228 GORRESPTMTFLCKECP-KIHIIHDIRIDKDVPEQEHMRMLHERFEIKDKMLIEFEY 286

245 -----CPSFDNVFGTDPSEVHIVRRVLLKEIPANAESSAMLMDSFKLKDLSDF-- 297

287 SPDEPKRRKPPGKSVNSKLSIKTL 311

298 ----NAQKRPNQRPEELSLVKCI 318

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

A:Residues: 1-374 <BRO>  
A:Cross-references: EMBL:Z29518; NID:g575959; PIDN:CAA82638.1; PID:g575960  
C:Superfamily: probable membrane protein YBR042c

Query Match 17.5%; Score 328; DB 2; Length 374;

Best Local Similarity 29.8%; Pred. No. 1.8e-20;

Matches 92; Conservative 64; Mismatches 121; Indels 32; Gaps 8;

18 VLMAGVRLISA-----FLPAR-----FYQALDRLCYVQSMVLEFFENTGYQ 62

9 VLPGLFLPLSLGLVNAIQAVLFTTIPRESKSFRRINRFLAELMLQVLWVVMAGVK 68

63 ILLYGD----LPRKKNIIYLANHOSTVDMIVADILAIRONALGHVRYLKEGLKMLPLY 118

69 VOLHADEYTRSMGKEHALIISNRSIDIMLIGMILAAQSGCGLSTLAVAKSKSFLPLV 128

119 G---CYFAOHGIIYKRSKFNKEMKRNKLOSVDAGTPMYLVIPPEGRTYNPEQTKVLSA 176

129 GWSMWFAYE--LFLERSMARKDEKTLKWLQRLKDFPRFWLALFVEGTRFPAK---LLA 183

177 SOAFPAORGAVLKLHVLPRIKATHVAFDCMKNYLDATYDVYVEGKDDGGRRESPTM 236

184 AQEYASQGLPARPNVLPRTKGFSAVSTMRDPVATYDITVVP-KDS-----PQPTM 237

237 TEFLCKECPKIHIIHDIRIDKDVPEQEHMRMLHERFEIKDKMLIEFESPPERRKRP 296

238 LRLKQSSYIHYRMKNHANSMPKSDSDVSKKCKDIFVAKDALLKHLATGTFDEEIRP 297

297 PGKSVNSKL 305

298 IGRPVKSL 306

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db





Db 122 -GAGMASSGSYIFLDNRFENDKPVLEIRIVKYSSEKKYOLLFAGETDKGERATRL- 178

QY 176 ASQAFARGLAVLKHVLTLPRIKATHVAFDCMK--NYLDAIVDYVYEGKDDGGQRRES 233

Db 179 -SDFAKNGKLPREYVILHRTTGFKFLMELMKENIKIYVDTITAYSG-----TIV 230

QY 234 PTMTFELCKECP-KIHIIHIDRIDKKDVPDEQEHMRMLHERFEIKDKMLIEFYSPD 289

Db 231 DTEAKLLAGNPPDKVHLDVKKYKRLDEIP-TGEGCEKMLTDLMATKERLKKFYEQEE 286

RESULT 8

T07936

probable glycerol-3-phosphate O-acyltransferase (EC 2.3.1.15) - rape

N:Alternate names: 1-acyl-sn-glycerol-3-phosphate acyltransferase

C:Species: Brassica napus (rape)

C>Date: 21-May-1999 #sequence\_revision 21-May-1999 #text\_change 20-Jun-2000

C:Accession: T07936

R:Brough, C.L.

submitted to the EMBL Data Library, June 1995

A:Reference number: 216230

A:Accession: T07936

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-311 <BRO>

A:Cross-references: EMBL:249860; NID:g1149594; PIDN:CAA90019.1

A:Experimental source: cv. jet neuf; embryo

C:Superfamily: probable membrane protein YBR042C

C:Keywords: acyltransferase; coenzyme A

Query Match 12.8%; Score 239; DB 2; Length 311;

Best Local Similarity 29.8%; Pred. No. 6,8e-13;

Matches 59; Conservative 46; Mismatches 79; Indels 14; Gaps 5;

QY 86 VDMIVADIIILRONALCHVRVYLKEGLKMLPLYG--CYFAOHGCIYVRSKRFNEKRN 143

Db 17 ISGLVNLQ-RSGCLGSALAVMKSSKFLPVIGSMWFSEY--LFLERMMWAKDESTIKS 73

QY 144 KLSYVDAGTPMYLVIPPEGTRVNEPTQKVLSSQAFARGLAVLKHVLTLPRIKATHVA 203

Db 74 GLORLNDPPRFMLALVEGTRFEAK---LKAQETAAASSELPVRNVLIPRKGVSA 130

QY 204 FDCMKNLDAIVDYVYEGKDDGGQRRESPTMTFELCKECPKIHIIHIDRIDKKDVEEQ 263

Db 131 VSNMRSPVPAIDMTVAIP-----KTSPPPTMLRLFKGQPSVYVHAIKCHSMKDEPSE 184

QY 264 EHMRLMLHERFEIKDKML 281

Db 185 DEIAOWCRODQFVTKDALL 202

RESULT 9

T06755

probable glycerol-3-phosphate O-acyltransferase (EC 2.3.1.15) - Arabidopsis thaliana

N:Alternate names: protein F15B8.160

C:Species: Arabidopsis thaliana (mouse-ear cress)

C>Date: 23-Apr-1999 #sequence\_revision 23-Apr-1999 #text\_change 19-May-2000

C:Accession: T06755

R:Queller, F.; Benes, V.; Rechmann, S.; Borkova, D.; Ansoerge, W.; Salanoubat, M.; Nemes, submitted to the Protein Sequence Database, April 1999

A:Reference number: 215794

A:Accession: T06755

A:Molecule type: DNA

A:Residues: 1-310 <QUE>

A:Cross-references: EMBL:AL049660; GSPDB:GNO0061; ATSP:F15B8.160

C:Experimental source: cultivar Columbia; BAC clone F15B8

C:Genetics:

A:Gene: ATSP:F15B8.160

A:Map position: 3

A:Intons: 26/3; 46/3; 72/3; 125/3; 167/3; 198/3; 227/3

C:Superfamily: probable membrane protein YBR042C

C:Keywords: acyltransferase; coenzyme A

Query Match 12.7%; Score 238.5; DB 2; Length 310;

Best Local Similarity 24.8%; Pred. No. 7.5e-13;

Matches 67; Conservative 64; Mismatches 106; Indels 33; Gaps 8;

QY 97 RONALGHVRVYLKEGLMPLYG--CYFAOHGCIYVRSKRFNEKRNKLSQYVDAGTP 154

Db 27 RSGCLGSALAVMKSSKFLPVIGSMWFSEY--LFLERMMWAKDESTIKSLGRLSDPRP 84

QY 155 MYLVIPPEGTRVNEPTQKVLSSQAFARGLAVLKHVLTLPRIKATHVADCKKNYDAI 214

Db 85 FWLALFVEGTRFEAK---LKAQETAAASSELPVRNVLIPRKGVSAVSNMRSFPAI 141

QY 215 YDVTYVEGKDDGGQRRESPTMTFELCKECPKIHIIHIDRIDKKDVPDEQEHMRMLHERF 274

Db 142 YDMIVTIP-----KTSPPPTMLRLFKGQPSVYVHAIKCHSMKMDLPESDAIAQWCHDQF 195

QY 275 EIKDKMLIEFYSPDPRRRRFPKGSVNSKLSIKTLPMLILSG-----LTAGML----- 325

Db 196 VAKDALLDKHIAA-----DTFPGQ--QEQNIGRPKSLAVVLSWACVLTGAIKFLHW 246

QY 326 --MTDAGRKLVTWITIGTLGCLMTYIKA 353

Db 247 AQLFSSWKGITISALGIIITLCMQILIRS 276

RESULT 10

S45900

probable membrane protein YBR042C - yeast (Saccharomyces cerevisiae)

N:Alternate names: hypothetical protein YBR0412

C:Species: Saccharomyces cerevisiae

C>Date: 26-Aug-1994 #sequence\_revision 09-Sep-1994 #text\_change 22-Oct-1999

C:Accession: S45900

R:Andre, B.; Cziepluch, C.; Hein, C.; Jauniaux, J.C.; Urrestiarazu, A.; Vissers, S. submitted to the Protein Sequence Database, August 1994

A:Reference number: S45903

A:Accession: S45900

A:Molecule type: DNA

A:Residues: 1-397 <AND>

A:Cross-references: EMBL:235911; NID:g536265; PIDN:CAA84984.1; PID:g536266; GSPDB:GNO

A:Experimental source: strain S288C

C:Genetics:

A:Gene: MIPS:YBR042C

A:Map position: 2R

C:Superfamily: probable membrane protein YBR042C

C:Keywords: transmembrane protein

F:12-37/Domain: transmembrane #status predicted <TM1>

F:55-77/Domain: transmembrane #status predicted <TM2>

F:134-150/Domain: transmembrane #status predicted <TM3>

F:372-390/Domain: transmembrane #status predicted <TM4>

Query Match 12.4%; Score 231.5; DB 2; Length 397;

Best Local Similarity 25.9%; Pred. No. 4.2e-12;

Matches 88; Conservative 55; Mismatches 136; Indels 61; Gaps 13;

QY 43 LYCYQSMVLEFFENTYGVGILIXDLPKN-----KENIYLANHGSTVDWIVADILAI 96

Db 67 LHVAVPAAVAVHTTENSVPKGTFFLDLKKRRLISHLKSNSVAACNQIYDWMIFLWMLAY 126

QY 97 RONALGHVRVYLKEGLKMLPLYGCFYFAOHGCIYVRSKRFNEKRNKL----- 145

Db 127 TSNLGNVFIILKKSILASTIILFGMKNYFIFMSKRMADKTTLSLSLAGLDSNARGAG 186

QY 146 -----QSYVDAGT-----PMVLVTPPEGTRVNEPTQKVLSSQAFQAQ 183

Db 187 SLAGKSPERITEEGESIMNDEVIDPKQIHWPYMLILPEEGTINSAD-TROKSAK--YAAK 243

QY 184 RGLAVLKHVLTLPRIKATHVAFDCMKNLDAIVDYVYEG--KDDGGQRRESPTMTFELC 241

Db 244 IGRKPKRNVLPHSTGLIRSLQIKPSIESLVDITIGYSGVQKEEGELIYG-LKSIIFLE 302

```

Oy 242 KECPEKIIHIDRIDKRVN-EEQEHMRMRMHERFEIKDKMLIEFYES----PPPERKKR 295
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 303 GKYPFLVDIHRARDVDKIDPLEDNEFESEWLKIKWSEKDALMERYSTGSEVSDPE---- 358
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Oy 296 FPGKSVNSKLSIKR-TLPSMLILSGTLFAGMLMTAGRRLY 334
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 359 -TNHSVYDSFKINRIELTEVLILPLTLIWLV----YKLY 393
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

RESULT 11
554641
probable membrane protein YDR018c - yeast (Saccharomyces cerevisiae)
N:Alternate names: hypothetical protein D3246; hypothetical protein p2f396; hypothetical
C:Species: Saccharomyces cerevisiae
C:Date: 08-Jul-1995 #sequence_revision 01-Sep-1995 #text_change 20-Jun-2000
C:Accession: S54641; S63425; S67831; S72116
R:Edman, K.; Brown, D.; Hamlyn, N.; Bowman, S.
submitted to the EMBL Data Library, May 1995
A:Reference number: S54638
A:Accession: S54641
A:Molecule type: DNA
A:Residues: 1-396 <DED>
A:Cross-references: EMBL:Z49770; NID:9840867; PIDN:CAA89843.1; PID:9840871
R:Experimental source: strain AB972
R:Eide, L.G.; Sander, C.; Prydz, H.
submitted to the EMBL Data Library, February 1996
A:Description: Sequencing and analysis of a 35.4 kb region on the left arm of chromosome
A:Reference number: S63416
A:Accession: S63425
A:Molecule type: DNA
A:Residues: 1-396 <EID>
A:Cross-references: EMBL:X95966; NID:91216215; PIDN:CAA65210.1; PID:91216225
R:Prydz, H.; Eide, L.G.
submitted to the Protein Sequence Database, July 1996
A:Reference number: S67822
A:Accession: S67831
A:Molecule type: DNA
A:Residues: 1-396 <PRY>
A:Cross-references: EMBL:Z7314; NID:91431443; PIDN:CAA98838.1; PID:91431444; MIPS:YDR01
A:Experimental source: strain S288C
R:Eide, L.G.; Sander, C.; Prydz, H.
Yeast 12, 1085-1090, 1996
A:Title: Sequencing and analysis of a 35.4 kb region on the left arm of chromosome IV fr
A:Reference number: S72107; MUID:97051598
A:Accession: S72116
A:Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-396 <EIB>
A:Cross-references: EMBL:X95966; NID:91216215; PIDN:CAA65210.1; PID:91216225
A:Note: the nucleotide sequence was submitted to the EMBL Data Library, February 1996
C:Genetics:
A:Map position: 4R
A:Note: YDR018c
C:Superfamily: probable membrane protein YBR042c
C:Keywords: transmembrane protein
F:27-43/Domain: transmembrane #status predicted <TM1>
F:69-85/Domain: transmembrane #status predicted <TM2>
F:376-392/Domain: transmembrane #status predicted <TM3>

Query Match 11.8%; Score 221.5; DB 2; Length 396;
Best Local Similarity 25.4%; Pred. No. 3.1e-11;
Matches 78; Conservative 49; Mismatches 113; Indels 67; Gaps 10;

Oy 73 KENIITLANHOSYVDWIYADILAIRONALGHVRYVYKEGLKMLPLYGCFYFHGGIYKR 132
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 108 KDRALIIINHOMQYMDWILMWLSEFVNSLGNVYIILKKALQIYPLILGGMNRFYFLSR 167
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Oy 133 SAKFEKEMRKRL-----GSYVAGTFPMY-LVPEPEGTRNPEQTK 172
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 168 NMQDKERKLTNSILVSMIDINARCKGGLTYNKCSCYSTKTNESIAIYNLIMPEBGTNIS---LK 224
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Oy 173 VLSASQAFPAARGTL-AVLKHLVLPRIKATHAVAFECMKNNYLDAIYADYVYVEGRDDGGOR 230

```

Db	225	TREKSEAFQPAHLDHVQLHLLPSPKSGIKRAVEKLADSLDAIVDTGY-----	275
Oy	231	RESPLM-----FEFLKEC-----PKIHIDRIDKQVP-DEQEHMRMLHEREPI	276
Db	276	--SFLATREYVGIKFTLTKLFLMGVYPEKVDPIREFRVNEIPLDDEVEFNWLLGWKE	333
Oy	277	KDKMLIEFESPDPERRRRPGKSVNSKLSIKT-----LPSMLILSGITAGML	325
Db	334	KDQLLEDYVNM-----GQKSNAKNDNOSIVVTTQTGFGHETLPLRLSYGGEFAPLI	387
Oy	326	MTDGRK 332	
Db	388	LVFVWKR 394	
RESULT	12		
	T25998		
	hypothetical protein ZK40.1 - Caenorhabditis elegans		
	C:Species: Caenorhabditis elegans		
	C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999		
	C:Accession: T25998		
	R:Bradshaw, H.		
	submitted to the EMBL Data Library, July 1996		
	A:Description: The sequence of C. elegans cosmid ZK40.		
	A:Reference number: Z20121		
	A:Accession: T25998		
	A:Status: preliminary; translated from GB/EMBL/DBJ		
	A:Molecule type: DNA		
	A:Residues: 1-523  		
	A:Cross-references: EMBL:U64839; PIDN:AAB04844.1; GSPDB:GNO0023; CESP:ZK40.		
	A:Experimental source: strain Bristol N2; clone ZK40		
	C:Genetics:		
	A:Gene: CESP:ZK40.1		
	A:Map position: 5		
	A:Introns: 7/1; 38/3; 93/3; 126/3; 160/1; 301/3; 422/2; 472/3		
Query Match	11.4%;	Score 213.5; DB 2; Length 523;	
Best Local Similarity	22.5%;	Pred No. 2.2e-10;	
Matches 80; Conservative 62; Mismatches 148; Indels 65; Gaps			
Oy	6	PSVLLGTAPTYVLAMGWMRLSAPLPAPFYQALDRLCYVQSNVLEFFENYTGVOILL	65
Db	167	PLVVL-----FLKPLRMQMDRLGVIMPGALCYIFGANIRI	208
Oy	66	YGDLPKNKENIYLIANKQSVDMIVADILAIKONALGH-----RYLKEGLKMWPL	117
Db	209	KGDIINHEPALILMNRITLDMV-----FFNNALYKKMDPMLCTEIKISLKGMLKYPG	262
Oy	118	YGCFAFHGGIYVYRSKAFKFEKEMRNKLOSVDAGTPMVLVFEFGSTRYNPEOTKVLAS	177
Db	263	AGMAMQAASTYFLDRSDTDTKTLNDILNLYAETEKYOLLFPBSTDKCRKATE--RS	319
Oy	178	QAFPAORGLAVLKHLVLRPIKATHVAFDCK--NYLDAIYDVTVVYEGKDDGGORRESPT	235
Db	320	RHSEKKGLVYGVVHLPRVTGVIYQAMRRANNIKYIYDVSIG-----GDAIYQSE	373
Oy	236	MTEFLCKECCR-THIHDIRDKKDQVEDEQEHMRMLHEFFELKDKMLIEFESPDPERK	294
Db	374	LDIFAHGVCPEKEVYGVIKYIPDAIEPQTDGALGOWLVNLMRKNEKRLKRYEM--PRNVR	431
Oy	295	RFPKGSVNSKLSIKTLPSMLILSGITAGMLMDGARKLYVNTWVLYGTLIGLAWY	349
Db	432	QFPDPGVEYEYLDNN-----TDRAQKGLIGFWCFITY--FWM	467
RESULT	13		
	T22689		
	hypothetical protein F55A11.5 - Caenorhabditis elegans		
	C:Species: Caenorhabditis elegans		
	C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999		
	C:Accession: T22689		



GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 28, 2002, 11:08:16 ; Search time 21.99 Seconds

(without alignments)  
621.555 Million cell updates/sec

Title: US-09-853-526-4

Perfect score: 1871

Sequence: 1 MRYLPSVVLGTAPRYVLA.....YVNTWVYGLGLMTVTRKA 353

Scoring table:

BIOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 105224 seqs, 38719550 residues

Total number of hits satisfying chosen parameters: 105224

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database : SwissProt\_40:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1871	100.0	353	1	PLCE_HUMAN
2	1568.5	83.8	354	1	PLCE_MOUSE
3	270	14.4	378	1	PLCD_RAT
4	268.5	14.4	378	1	PLCD_HUMAN
5	240	12.8	376	1	PLCC_HUMAN
6	231.5	12.4	397	1	YBP2_YEAST
7	221.5	11.8	396	1	Y205_YEAST
8	204.5	10.9	370	1	Y181_YEAST
9	182	9.7	391	1	Y173_CAEL
10	154	8.2	310	1	Y181_CAEL
11	120.5	6.4	303	1	PLSC_YEAST
12	119	6.4	281	1	PLSC_LIMDO
13	116	6.2	281	1	PLSC_LIML
14	114	6.1	266	1	PLSC_MYCPN
15	111.5	6.0	488	1	EXON_HSV6
16	111	5.9	377	1	VENY_FOMPV
17	109.5	5.9	488	1	EXON_HSV6
18	109	5.8	240	1	PLSC_HAELN
19	104	5.6	240	1	PLSC_HELPI
20	104	5.6	308	1	PLSC_COCPN
21	101	5.4	262	1	PLC1_CAEL
22	99.5	5.3	237	1	PLSC_HELPI
23	99	5.3	245	1	PLSC_SALTY
24	99	5.3	1551	1	YQ12_CAEL
25	97	5.2	245	1	PLSC_ECOLI
26	96	5.1	484	1	Y142_NPVP
27	94.5	5.1	282	1	PLC2_CAEL
28	93	5.0	280	1	PLSC_BORBU
29	90	4.8	892	1	RA16_SCHPO
30	88.5	4.7	446	1	MY64_MOUSE
31	86	4.6	493	1	SLK_ARCFU
32	85.5	4.6	160	1	ALGO_PSEAR
33	85	4.5	561	1	SNTC_HUMAN

34	84.5	4.5	255	1	PLSC_NEIMB	Q91247 neisseria m
35	84	4.5	1272	1	YZ28_METJA	O60287 methanococ
36	83.5	4.5	255	1	PLSC_METJO	O59601 neisseria g
37	83.5	4.5	255	1	PLSC_NEIMA	O91041 neisseria m
38	83.5	4.5	268	1	PLSC_MYCSE	O49402 mycoplasma
39	83.5	4.5	2386	1	RAD3_SCHPO	O02099 schizosach
40	83	4.4	318	1	YOK3_CAEL	O09287 caenorhabd
41	83	4.4	778	1	YF05_METJA	O58900 methanococ
42	83	4.4	1241	1	TRK1_SACRA	P28569 saccharomy
43	82.5	4.4	329	1	SDB1_ECOLI	P06997 escherichia
44	82.5	4.4	998	1	PPOL_XENLA	P31669 xenopus lae
45	82	4.4	386	1	YMK7_YEAST	Q03760 saccharomy

## ALIGNMENTS

RESULT ID	PLCE_HUMAN	STANDARD:	PRT:	353 AA.
AC	Q9N0Q2; Q9N0G4;			
DT	01-MAR-2002 (Rel. 41, Created)			
DT	01-MAR-2002 (Rel. 41, Last sequence update)			
DT	01-MAR-2002 (Rel. 41, Last annotation update)			
DE	1-acyl-sn-glycerol-3-phosphate acyltransferase epsilon (EC 2.3.1.51)			
DE	(1-AGP acyltransferase 5) (1-AGPAT 5) (lysophosphatidic acid acyltransferase-epsilon) (LPAAT-epsilon) (1-acylglycerol-3-phosphate O-acyltransferase 5).			
DE	O-acyltransferase 5).			
GN	AGPAT5.			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.			
OX	NCBI_TaxID=9606;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RA	Leung D.W.;			
RT	*Cloning and expression of LPAAT-epsilon.*;			
RL	Submitted (MAY-2001) to the EMBL/Genbank/DBJ databases.			
RN	[2]			
RP	SEQUENCE FROM N.A.			
RA	Cohen D., Chumakov I., Blumenfeld M., Bougueleret L.;			
RL	Patent number WO932644, 01-JUL-1999.			
RN	[3]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=Amgdala;			
RX	MEDLINE=21154917; PubMed=11230166;			
RA	Medline S., Well B., Wellenreuther R., Gassenhuber J., Glassl S.,			
RA	Ansoerge W., Boecker H., Bloeker H., Bauersachs S., Blum H.,			
RA	Lauber J., Duesterhoeft A., Beyer A., Koehrer K., Strack N.,			
RA	Mewes H.-W., Oltmannseder B., Obermaier B., Tampe J., Heubner D.,			
RA	Wamblut R., Korn B., Klein M., Poustka A.;			
RT	*Towards a catalog of human genes and proteins: sequencing and			
RT	analysis of 500 novel complete protein coding human cDNAs.*;			
RN	Genome Res. 11:422-435(2001).			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=Placenta;			
RA	Isogai T., Ota T., Hayashi K., Sugiyama T., Otsuki T., Suzuki Y.,			
RA	Nishikawa T., Nagai K., Sugano S., Shiratori A., Sudo H.,			
RA	Wagatsuna M., Hosokiri T., Kaku Y., Kodaira H., Kondo H., Sugawara M.,			
RA	Takahashi M., Chiba Y., Ishida S., Murakawa K., Ono Y., Takiguchi S.,			
RA	Matanabe S., Kimura K., Murakami K., Ishi S., Kawai Y., Saito K.,			
RA	Yamamoto Y., Wakamatsu A., Nakamura Y., Nagahara K., Masuno Y.,			
RA	Nimomiya K., Iwayanagi T.;			
RT	*NEO human cDNA sequencing project.*;			
RT	Submitted (FEB-2000) to the EMBL/Genbank/DBJ databases.			
CC	-1- FUNCTION: CONVERTS LYOPHOSPHATIDIC ACID (LPA) INTO PHOSPHATIDIC			
CC	ACID BY INCORPORATING ACYL MOIETY AT THE 2 POSITION (BY			
CC	SIMILARITY).			
CC	-1- CATALYTIC ACTIVITY: Acyl-CoA + 1-acyl-sn-glycerol 3-phosphate -			
CC	CoA + 1,2-diacyl-sn-glycerol 3-phosphate.			
CC	-1- PATHWAY: SECOND STEP IN DE NOVO PHOSPHOLIPID BIOSYNTHESIS.			
CC	-1- SUBCELLULAR LOCATION: Integral membrane protein (potential).			

CC -1- SIMILARITY: BELONGS TO THE 1-ACYL-SN-GLYCEROL-3-PHOSPHATE  
 CC ACYLTRANSFERASE FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL: AF375789; AAK54809.1; ALT\_INT.  
 DR EMBL: AL136587; CAB66522.1; ALT\_INT.  
 DR EMBL: AK002072; BAA92069.1; -  
 DR InterPro: IPR002123; Acyltransferase.  
 DR Pfam: PF01553; Acyltransferase; 1.  
 KM Phospholipid biosynthesis; Transferase; Acyltransferase;  
 KM Transmembrane.  
 FT TRANSMEM 7 29 POTENTIAL.  
 FT TRANSMEM 44 66 POTENTIAL.  
 FT TRANSMEM 334 351 POTENTIAL.  
 FT CONFLICT 145 145 L -> V (IN REF. 2).  
 SO SEQUENCE 353 AA; 40813 MW; A05B1FA246CE1B64 CRC64;

Query Match 100.0%; Score 1871; DB 1; Length 353;  
 Best Local Similarity 100.0%; Pred. No. 1.1e-152;  
 Matches 353; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MRYLLPSVLLGTPPYVLAWGVWRLLSAFLPARFYQALDRLCYQYQSVLFEFFENYTG 60  
 DB 1 MRYLLPSVLLGTPPYVLAWGVWRLLSAFLPARFYQALDRLCYQYQSVLFEFFENYTG 60  
 QY 61 VOILLGDLPRKKNENIYYLANHSTYDVIADILAIROMLGHVRYLKGKMLPLCYG 120  
 DB 61 VOILLGDLPRKKNENIYYLANHSTYDVIADILAIROMLGHVRYLKGKMLPLCYG 120  
 QY 121 YFAOHGGIYVRSKAFNEKEMRNKLQSYVDAGTPMYLVFPESTRYNEPQTKVLSAQAF 180  
 DB 121 YFAOHGGIYVRSKAFNEKEMRNKLQSYVDAGTPMYLVFPESTRYNEPQTKVLSAQAF 180  
 QY 181 AAOAGLAVLKHVLTLPRIKATHVAFDCKKNYLDIYDVTYVEGKDDGGORRESPTMTEFL 240  
 DB 181 AAOAGLAVLKHVLTLPRIKATHVAFDCKKNYLDIYDVTYVEGKDDGGORRESPTMTEFL 240  
 QY 241 CKCEPKIHIDRIDKDDVPEDEQEHMRWMLHEFEIKDKMLIEFYESPDERRRKRPFGKS 300  
 DB 241 CKCEPKIHIDRIDKDDVPEDEQEHMRWMLHEFEIKDKMLIEFYESPDERRRKRPFGKS 300  
 QY 301 VNSKLSIKKTLPSMLISGLTAGMLMTDAGRKLYVNTWITGTLGCLMTVITKA 353  
 DB 301 VNSKLSIKKTLPSMLISGLTAGMLMTDAGRKLYVNTWITGTLGCLMTVITKA 353  
 RESULT 2  
 PILE\_MOUSE  
 ID PILE\_MOUSE STANDARD; PRT; 354 AA.  
 AC 09D1E8;  
 DT 01-MAR-2002 (Rel. 41, Last sequence update)  
 DT 01-MAR-2002 (Rel. 41, Last annotation update)  
 DE 1-acyl-sn-glycerol-3-phosphate acyltransferase epsilon (EC 2.3.1.51)  
 DE (1-AGPAT acyltransferase 5) (1-AGPAT 5) (Lysophosphatidic acid  
 DE acyltransferase-epsilon) (LPAAT-epsilon) (1-acylglycerol-3-phosphate  
 DE 0-acyltransferase 5).  
 GN AGPAT5.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniala; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathu; Muridae; Mus.  
 OX NCBI\_Taxid=10090;  
 RP [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=C57BL/6J; TISSUE=Embryo;

RX MEDLINE=21085660; PubMed=11217851;  
 RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,  
 RA Arakawa T., Hara A., Fukunishi Y., Kono H., Adachi J., Fukuda S.,  
 RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamana I.,  
 RA Saito T., Okazaki Y., Gojohori T., Bono H., Kasukawa T., Saito R.,  
 RA Kadoya K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,  
 RA Fleischmann W., Gaasterland T., Glass C., King B., Kochia H.,  
 RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,  
 RA Sakurai L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,  
 RA Sakai K., Okido T., Furuno M., Aono H., Balderelli R., Barsh G.,  
 RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,  
 RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,  
 RA Gostungsten S., Hill D., Hornann M., Hume D.A., Kamiya M., Lee N.H.,  
 RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,  
 RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,  
 RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,  
 RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whitaker C., Wilming L.,  
 RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawai H., Kohsaki S.,  
 RA Hayashizaki Y.;  
 RA "Functional annotation of a full-length mouse cDNA collection."  
 RL Nature 409:685-690(2001).  
 CC -1- FUNCTION: CONVERTS LYSOEPHOSPHATIDIC ACID (LPA) INTO PHOSPHATIDIC  
 CC ACID BY INCORPORATING ACYL MOIETY AT THE 2 POSITION (BY  
 CC SIMILARITY).  
 CC -1- CATALYTIC ACTIVITY: Acyl-CoA + 1-acyl-sn-glycerol 3-phosphate -  
 CC CoA + 1,2-diacyl-sn-glycerol 3-phosphate.  
 CC -1- PATHWAY: SECOND STEP IN DE NOVO PHOSPHOLIPID BIOSYNTHESIS.  
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein (potential).  
 CC -1- SIMILARITY: BELONGS TO THE 1-ACYL-SN-GLYCEROL-3-PHOSPHATE  
 CC ACYLTRANSFERASE FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL: AK003649; BAB22915.1; -  
 DR MG: MG1:1915880; Agpat5.  
 DR InterPro: IPR002123; Acyltransferase.  
 DR Pfam: PF01553; Acyltransferase; 1.  
 KW Phospholipid biosynthesis; Transferase; Acyltransferase;  
 KM Transmembrane.  
 FT TRANSMEM 7 29 POTENTIAL.  
 FT TRANSMEM 44 66 POTENTIAL.  
 FT TRANSMEM 335 352 POTENTIAL.  
 SO SEQUENCE 354 AA; 40943 MW; 447BE924B91E800 CRC64;

Query Match 83.8%; Score 1568.5; DB 1; Length 354;  
 Best Local Similarity 81.4%; Pred. No. 8.1e-127;  
 Matches 288; Conservative 39; Mismatches 26; Indels 1; Gaps 1;

QY 1 MRYLLPSVLLGTPPYVLAWGVWRLLSAFLPARFYQALDRLCYQYQSVLFEFFENYTG 60  
 DB 1 MRYLLPSVLLGTPPYVLAWGVWRLLSAFLPARFYQALDRLCYQYQSVLFEFFENYTG 60  
 QY 61 VOILLGDLPRKKNENIYYLANHSTYDVIADILAIROMLGHVRYLKGKMLPLCYG 120  
 DB 61 VOILLGDLPRKKNENIYYLANHSTYDVIADILAIROMLGHVRYLKGKMLPLCYG 120  
 QY 121 YFAOHGGIYVRSKAFNEKEMRNKLQSYVDAGTPMYLVFPESTRYNEPQTKVLSAQAF 180  
 DB 121 YFAOHGGIYVRSKAFNEKEMRNKLQSYVDAGTPMYLVFPESTRYNEPQTKVLSAQAF 180  
 QY 181 AAOAGLAVLKHVLTLPRIKATHVAFDCKKNYLDIYDVTYVEGKDDG-QORRESPTMTEFL 239  
 DB 181 AAOAGLAVLKHVLTLPRIKATHVAFDCKKNYLDIYDVTYVEGKDDG-QORRESPTMTEFL 239  
 QY 240 LKCEPKIHIDRIDKDDVPEDEQEHMRWMLHEFEIKDKMLIEFYESPDERRRKRPFGK 299  
 DB 240 LKCEPKIHIDRIDKDDVPEDEQEHMRWMLHEFEIKDKMLIEFYESPDERRRKRPFGK 299





```

Db 133 GMMWFTET--VFCSKRWQDRTVATSTSDHCRDYEPKXFFLIHCSTFTEKKHEI--- 187
Oy 177 SOAFAQRLAVLKHVLPRIKATVAFDCMKNYDAIDVTVVYEGKDDGGORRSPM 236
Db 188 SMOVARAKLPLRKHLLPRTKGFATVRSLSRVSAVVDCLNF-----RNENPIL 240
Oy 237 TEFLEKCPKIHIDRIDKKDVEQEHMRMLHERFEIKDKMLIEFY-----ESP-D 289
Db 241 LGVLNCKKHYADLYVRIPLEDIPEDDECSAMLHKLYQEKDAFOEYRTGTFPETPMV 300
Oy 290 PERRRFPKSVN 302
Db 301 PPRR---PWTLVN 310

RESULT 5
PICC_HUMAN STANDARD: PRT: 376 AA.
AC Q9NRZ7; Q9NRZ6;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE 1-acyl-sn-glycerol-3-phosphate acyltransferase gamma (EC 2.3.1.51) (1-AGP acyltransferase 3) (1-AGPAT 3) (lysophosphatidic acid acyltransferase-gamma) (LPAT-gamma) (1-acylglycerol-3-phosphate O-acyltransferase 3).
DE GN AGPAT3.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A. (ISOFORMS GAMMA-1 AND GAMMA-2).
RA Leung D.W.;
RT "Structure and functions of lysophosphatidic acid acyltransferases.";
RL Submitted (JUN-1999) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Fetal liver;
RA Nagamine K., Kudoh J., Minoshima S., Kawasaki K., Hase T., Shimizu N.;
RT "Isolation of a novel gene encoding 1-acylglycerol-3-phosphate O-acyltransferase 3 (AGPAT3) from the human chromosome 21q22.3.";
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RA Strausberg R.;
RT Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: CONVERTS LYOPHOSPHATIDIC ACID (LPA) INTO PHOSPHATIDIC ACID BY INCORPORATING ACYL MOIETY AT THE 2 POSITION (BY SIMILARITY).
CC -1- CATALYTIC ACTIVITY: Acyl-CoA + 1-acyl-sn-glycerol 3-phosphate = CoA + 1,2-diacyl-sn-glycerol 3-phosphate.
CC -1- PATHWAY: SECOND STEP IN DE NOVO PHOSPHOLIPID BIOSYNTHESIS.
CC -1- SUBCELLULAR LOCATION: Integral membrane protein (Potential).
CC -1- ALTERNATIVE PRODUCTS: 2 ISOFORMS: GAMMA-1 (SHOWN HERE) AND GAMMA-2: ARE PRODUCED BY ALTERNATIVE SPLICING.
CC -1- SIMILARITY: BELONGS TO THE 1-ACYL-SN-GLYCEROL-3-PHOSPHATE ACYLTRANSFERASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AF156774; AAF80336.1; -
DR EMBL: AF156775; AAF80337.1; -
DR EMBL: AB040138; BAB18943.1; -

```

```

DR EMBL: BC011971; AAH11971.1; -
DR InterPro: IPR002123; Acyltransferase.
DR Pfam: PF01553; Acyltransferase; 1.
KW Phospholipid biosynthesis; Transferase; Acyltransferase;
KW Transmembrane; Alternative splicing.
FT TRANSMEM 15 39 POTENTIAL.
FT TRANSMEM 126 144 POTENTIAL.
FT TRANSMEM 308 329 POTENTIAL.
FT TRANSMEM 336 352 POTENTIAL.
FT VARSPIC 1 62 MISSING (IN ISOFORM GAMMA-2).
SQ SEQUENCE 376 AA; 43381 MW; C12CDBB7CC363852 CRC64;

Query Match 12.8%; Score 240; DB 1; Length 376;
Best Local Similarity 25.4%; Pred. No. 4e-13;
Matches 69; Conservative 63; Mismatches 116; Indels 24; Gaps 7;

Oy 34 RPYQALDDRLCYQSMVLFEEFENTGVQILKGLPK-----NKENIYILANHOSTVMI 89
Db 44 QLYRLNCRSLAYSLMSQLVLMWMSCTECTLTQATYVERFGKEHAYIILHNFEIDPL 103
Oy 90 VADIIAIPRONALGHVRYVLKEGLKMLPLYG--CYPAQHGGIYKRSAPFNEKEMNKLOS 147
Db 104 CGWTMGERRGVLSGSKVLAKEKELLYPLGLGWTWYFLEI--VECKRKWEDRQTVVGLLR 161
Oy 148 YVDAGTPMYLVIPFECTRYNPQOTKYVLSQAFAQRLAVLKHVLPRIKATVAFDCM 207
Db 162 LSDPYRYMFWFLLYCEGTREPTETKHHV---SMEVAAAKGLPVLKHLIDPRTKGFTTAVVCL 218
Oy 208 KNYLDAIVDVTVVYBEKDKDGGRRSPMTREFLKCEPRKIHIDRIDKKDVEQEHMR 267
Db 219 RGTVAADVTVTLNFRG-----NKNPSLGLIGLYKTYEADKCVRRFPLEDIPDEKEA 271
Oy 268 RMLHERFEIKDKMLIEFYESPDPERRKRPCK 299
Db 272 QMLHKRYQEKD-ALQETI-----NKGMPGE 297

RESULT 6
YBP2_YEAST STANDARD: PRT: 397 AA.
AC P38226;
DT 01-OCT-1994 (Rel. 30, Created)
DT 01-OCT-1994 (Rel. 30, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hypothetical 45.5 kDa protein in FAY1-TCM62 intergenic region.
GN YBR042C OR YBR0412.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes; Saccharomycetales; Saccharomycetaceae; Saccharomyces.
OX NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=S288C;
RA Andre B., Cziepluch C., Hein C., Jauniaux J.-C., Urrastarazu A., Vissers S.;
RT Submitted (AUG-1994) to the EMBL/GenBank/DBJ databases.
CC -1- SUBCELLULAR LOCATION: Integral membrane protein (Potential).
CC -1- SIMILARITY: BELONGS TO THE 1-ACYL-SN-GLYCEROL-3-PHOSPHATE ACYLTRANSFERASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).
CC -----
DR EMBL: Z35911; CAA84984.1; -
DR PIR: S45900; S45900.
DR SGD: S0000246; YBR042C.
DR InterPro: IPR002123; Acyltransferase.

```





CC -1- SIMILARITY: BELONGS TO THE 1-ACYL-SN-GLYCEROL-3-PHOSPHATE  
CC ACYLTRANSFERASE FAMILY.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
CC EMBL: D86960: BAA13196.1: -  
CC DR InterPro: IPR002123: Acyltransferase.  
CC KW Hypothetical protein; Phospholipid biosynthesis; Transferase;  
CC ACYLTRANSFERASE; Transmembrane.  
CC FT TRANSMEM 22 42 POTENTIAL.  
CC FT TRANSMEM 342 362 POTENTIAL.  
CC SQ SEQUENCE 370 AA: 43089 MW: 93498544EA651541 CRC64:

Query Match 10.9%; Score 204.5; DB 1; Length 370;  
Best Local Similarity 22.8%; Pred. No. 4.3e-10;  
Matches 90; Conservative 66; Mismatches 143; Indels 95; Gaps 18;

OY 1 MRYLLPSVVLCTAPRYVLAMGVMLLSAFAPRYQALDRLCYVQSMVLEFFENYTG 60  
DB 20 MRFAPVYVNNLVAPISYCYVILLOPLRLDSKRFMY-IEGIMYKMLGMVA-SMGWYAG 77  
OY 61 VOILLYGDLPR-KNENIITYLANHOSTVDIVADILAIQN---ALGHRYVYAKELKML 115  
DB 78 YVVMGEGEDIKAVSKDEAVLVNHOATGD-VCTLMMLQDDKGLVVAQMMMLDHIKKT 135  
OY 116 PLVYCGFAHGIGYVRSKAFENE---NRNKLSYVDAGTMYLYIFEGFTYNEQTR 172  
DB 136 N-FGILSVLHGDFEIRQGSYRDQOLLKHLHENYRSDRKMYLLEFGGGLR---K 190  
OY 173 VLSASQAFQAQGLAVLKHVLPRIKATVAFDCM---KN-----YLDA----- 213  
DB 191 RRETSGAFKAKNNLPFLVPLPRSGATKIILNALVAQOKNSPAGDAKELDSKSGIQ 250  
OY 214 -LYDVYVVEGRKD-----GGORESPMTPEFLCKECPKIHIDRIDKDKVPEOE 265  
DB 251 WIDDTTIAPKAEPIDIOFWILGYRK--PTVT-----HMYRIFPKIDVPLETDD 298  
OY 266 MRRWLHERFEIKDKMLIEFES---PDPERRRKFPKSVNSKLSIKTLPSMLILSLGTA 322  
DB 299 LRTWLYQRVEKEDLSHYETGAFP-----PSKGKKAVSREMTLS----- 340  
OY 323 GMLMTDAGRKLYVNTWI---YGTLLGCLWVTI 351  
DB 341 -----NLWIFLIQSFAFLSGYMWYNI 361

RESULT 9  
Y773-CAEEL STANDARD: PRT; 391 AA.  
AC Q11087;  
DT 01-NOV-1997 (Rel. 35, Created)  
DT 01-NOV-1997 (Rel. 35, Last sequence update)  
DT 01-NOV-1997 (Rel. 35, Last annotation update)  
DE Hypothetical 44.7 kDa protein COIC10.3 in chromosome X.  
GN COIC10.3.  
OS Caenorhabditis elegans.  
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;  
OC Rhabditidae; Pelodermidae; Caenorhabditis.  
OX NCBI\_Taxid=6239;  
RN [1]  
RN SEQUENCE FROM N.A.  
RC STRAIN=BRISTOL N2;  
RA Pauley A.;  
RL Submitted (MAR-1995) to the EMBL/GenBank/DBJ databases.  
CC -1- SIMILARITY: SOME. TO C.ELEGANS F08G5.2.  
CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
CC EMBL: U23526: AAB36850.1: -  
CC DR WormPep: COIC10.3: CE02449.  
CC DR InterPro: IPR002123: Acyltransferase.  
CC KW Hypothetical protein.  
CC SQ SEQUENCE 391 AA: 44718 MW: FB5C062A0172A8C1 CRC64:

Query Match 9.7%; Score 182; DB 1; Length 391;  
Best Local Similarity 25.0%; Pred. No. 3.9e-08;  
Matches 72; Conservative 47; Mismatches 105; Indels 64; Gaps 15;

OY 58 YTGVOILLYGD--LPKNENIITYLANHOSTVDIVADILAIROMALGHVR---YVLKEG 111  
DB 98 FVGATVTEYGTNLGYAEERKCLLANHGLLDHFV---LMQSLNGKGSIRSMWYIYNI 154  
OY 112 LKWLPLYGCTFAHOGGITY---KSAKFN--EKEMRKLOSIVDAGTPTLVIFPEGT 164  
DB 155 WKYTP-LGYMWTSHGNFVNGGVSKRDSVLSSEFRDLKNSFYKY-DYG---WVIMYPEGS 209  
OY 165 RYNPEQTKVLSASQAFQAQGLAVLKHVLPRIKATVAFD-----CMKNY- 210  
DB 210 RL-----YLVKNSGRTPFAENGLKPLDNCYPRPGAHAVALDVLCPPDLSLSKSGKGP 265  
OY 211 LDAIVDVYVVEGRKDDGGORRESP---TWTPEFLCKECPKIHIDRIDKDKVPEOE 265  
DB 266 IKYIIDATIGY-----RKGAVPIDICVDMGDMESVEASQPAVHYDVIPKPEMSDNL 318  
OY 266 MRRWLHERFEIKDKMLIEFESPDPERRRKFPKSVNSKLSIKTLPS 313  
DB 319 LKEFLERYIIDKDLAEFYKT-----GHFPGDRT---KVIPN 353

RESULT 10  
YIHG-ECOLI STANDARD: PRT; 310 AA.  
AC P32129;  
DT 01-OCT-1993 (Rel. 27, Created)  
DT 01-OCT-1993 (Rel. 27, Last sequence update)  
DT 16-OCT-2001 (Rel. 40, Last annotation update)  
DE Hypothetical protein yihg.  
GN YIHG OR B3862.  
OS Escherichia coli.  
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;  
OC Escherichia.  
OX NCBI\_Taxid=562;  
RN [1]  
RN SEQUENCE FROM N.A.  
RC STRAIN=K12 / MG1655;  
RX MEDLINE=93347969; PubMed=8346018;  
RA Plunkett G., Iii, Burland V.D., Daniels D.L., Blattner F.R.;  
RT "Analysis of the Escherichia coli genome. III. DNA sequence of the  
RT region from 87.2 to 89.2 minutes."  
RL Nucleic Acids Res. 21:3591-3598(1993).  
CC -1- SIMILARITY: BELONGS TO THE 1-ACYL-SN-GLYCEROL-3-PHOSPHATE  
CC ACYLTRANSFERASE FAMILY.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
CC EMBL: L19201; AAB02997.1; -

DR EMBL: AE000461; AAC76860.1; -.  
 DR PIR: S40808; S40808.  
 DR Eccegene: EG11833; yihg.  
 DR InterPro: IPR002123; Acyltransferase.  
 DR Pfam: PF01553; Acyltransferase; 1.  
 KW Hypothetical protein; Complete proteome.  
 SQ SEQUENCE 310 AA; 36289 MW; 9F8E3F52EB0B186E CRC64;

Query Match 8.2%; Score 154; DB 1; Length 310;  
 Best Local Similarity 20.6%; Pred. No. 7.2e-06;  
 Matches 64; Conservative 64; Mismatches 111; Indels 72; Gaps 14;

OY 1 MRVLLPSVVLGTAPTVYLVAMGWRLSAFLPAREFOALDDRLCYQSMVFEFFENYTG 60  
 DB 43 VLLLPVPV-----IMRKVSF-----CDFMYKCEGLAVLHLHN-PH 80  
 OY 61 VOILLYGLDPKKNEN-IYLANHOSVDWIVADILAIKONALGHV--RYVLKEGLKWL 116  
 DB 81 LQWEVHGLEGLSKRWYLLICNHRSMADIVLCVLFERK-----HLPNKYFLKQLAWVP 135  
 OY 117 LYG--C-----YFNQHGSIYKRSKAFKEMKRNKLSQSYVDACT-PMVLVIFPESTRN 167  
 DB 136 FLGLACWSLIDMPKPKRYSRVALRHPERRGKDVETTRRSCERFLHPTIIVNVEGSRPT 195  
 OY 168 PEOTFVLSASQAFAAQROGLAVLKHVLTPIKATHYAFDCMKVYLDIYDVYVYEGKDDG 227  
 DB 196 QEK-----HQQHTSTFQNLIPKKAAGIAMALNVLGKQFDKLANVTLCT--PDNN 242  
 OY 228 GORRESPTWTEFLCKECPKIHIDRIDKKDVEE-----QEHMRRLHERFEI 276  
 DB 243 RQ-----PFEDMLSGKRLRIVH---VDLQPIADELHGDYINDKSPRHRFQQLMSLMQ 294  
 OY 277 KDKMLIEFYES 287  
 DB 295 KDLRLTSLMSS 305

RESULT 11  
 PLSC\_YEAST STANDARD: PRT; 303 AA.

AC P33333;  
 DT 01-FEB-1994 (Rel. 28, Created)  
 DT 01-FEB-1994 (Rel. 28, Last sequence update)  
 DT 15-DEC-1998 (Rel. 37, Last annotation update)  
 DE Probable 1-acyl-sn-glycerol-3-phosphate acyltransferase (EC 2.3.1.51)  
 DE (1-AGP acyltransferase) (1-AGPAT) (Lysophosphatidic acid  
 acyltransferase) (LPAAT).  
 GN SLG1 OR YDL052C.  
 OS Saccharomyces cerevisiae (Baker's yeast).  
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;  
 OC Saccharomycetales; Saccharomycetaceae; Saccharomycos.  
 NC NCB1\_TaxID=4932;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=94012814; PubMed=8408076;  
 RA Naele M.M., Wells G.B., Lester R.L., Dickson R.C.;  
 RT "A suppressor gene that enables Saccharomyces cerevisiae to grow  
 without making sphingolipids encodes a protein that resembles an  
 Escherichia coli fatty acyltransferase";  
 RL J. Biol. Chem. 268:22156-22163(1993).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RX urrestarazu L.A., Andre B., Vissers S.;  
 RL Submitted (JUL-1996) to the EMBL/Genbank/DBJ databases.  
 RN [3]  
 RP SEQUENCE OF 182-303 FROM N.A.  
 RA Bloeker H., Brandt P.;  
 RL Submitted (JUL-1996) to the EMBL/Genbank/DBJ databases.  
 CC -1- FUNCTION: MAY BE AN ACYLTRANSFERASE WITH AN ALTERED SUBSTRATE  
 CC SPECIFICITY THAT ENABLES IT TO USE A C-26-COA IN PLACE OF THE  
 CC C-16 OR C-18-COAS USED BY THE WILD TYPE PROTEIN.  
 CC -1- CATALYTIC ACTIVITY: Acyl-CoA + 1-acyl-sn-glycerol 3-phosphate =

COA + 1,2-diacyl-sn-glycerol 3-phosphate.  
 -1- PATHWAY: FIRST STEP IN DE NOVO PHOSPHOLIPID BIOSYNTHESIS.  
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein (potential).  
 CC -1- SIMILARITY: BELONGS TO THE 1-ACYL-SN-GLYCEROL-3-PHOSPHATE  
 CC ACYLTRANSFERASE FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

DR EMBL: L13282; AAA16514.1; -.  
 DR EMBL: Z74100; CA98614.1; -.  
 DR PIR: A48600; A48600.  
 DR SGD: S0002210; SLG1.  
 DR InterPro: IPR002123; Acyltransferase.  
 DR Pfam: PF01553; Acyltransferase; 1.  
 KW Phospholipid biosynthesis; Transferase; Acyltransferase;  
 KW Transmembrane.  
 FT TRANSMEM 16 30 POTENTIAL.  
 FT VARIANT 44 44 Q -> L (IN SUPPRESSOR ALLELE SLG1-1).  
 SQ SEQUENCE 303 AA; 33887 MW; 36CBBC2659655EB CRC64;

Query Match 6.4%; Score 120.5; DB 1; Length 303;  
 Best Local Similarity 29.7%; Pred. No. 0.0051;  
 Matches 44; Conservative 20; Mismatches 61; Indels 23; Gaps 6;

OY 45 CYQSMVLEFFENYTVQILLYGLDLPKKNENIYLANHOSVDWIVADILAIKONALGHV 104  
 DB 50 CFYHVKML-----GLDKVYGEENLAKPKYIMINHOSITDIEP-----LGRI 94  
 OY 105 -----RYVLKEGLKWLPLYGCFYFAQHGSIYKRSKAFKEMKRNKLSQSYVDACTPMYTVI 159  
 DB 95 FPGGCTVTAKKSLKYVPFLGWFMAISGTYFLDRSKROEALIDLKGLENVKK-NKRALMV 153  
 OY 160 PPEGTR-VNPEOTKVLASQAF-AAORG 185  
 DB 154 PPEGTRSYNSELTMLEPFKKGAFLAOG 181

RESULT 12  
 PLSC\_LIMDO STANDARD: PRT; 281 AA.

AC Q42870; Q40120;  
 DT 15-DEC-1998 (Rel. 37, Created)  
 DT 15-DEC-1998 (Rel. 37, Last sequence update)  
 DT 15-DEC-1998 (Rel. 37, Last annotation update)  
 DE 1-acyl-sn-glycerol-3-phosphate acyltransferase (EC 2.3.1.51) (1-AGP  
 acyltransferase) (1-AGPAT) (Lysophosphatidic acid acyltransferase)  
 DE (LPAAT).  
 GN PLSC.  
 OS Limnantes douglasii (Douglas's meadowfoam).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
 OC eurosids II; Brassicales; Limnanthaceae; Limnantes.  
 NC NCB1\_TaxID=28973;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=96028122; PubMed=7588719;  
 RA Hanke C., Wolter F.P., Coleman J., Peterek G., Frentzen M.;  
 RT "A plant acyltransferase involved in triacylglycerol biosynthesis  
 RT complements an Escherichia coli sn-1-acylglycerol-3-phosphate  
 RT acyltransferase mutant";  
 RL Eur. J. Biochem. 232:806-810(1995).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=96046746; PubMed=7579178;  
 RA Brown A.P., Brough C.L., Kroon J., Slabas A.R.;  
 RT "Identification of a cDNA that encodes a 1-acyl-sn-glycerol-3-

RT phosphate acyltransferase from Limnantes douglasii.  
 RL Plant Mol. Biol. 29:267-278(1995).  
 CC -1- FUNCTION: CONVERTS LYOPHOSPHATIDIC ACID (LPA) INTO PHOSPHATIDIC  
 CC ACID BY INCORPORATING ACYL MOIETY AT THE 2 POSITION. THIS ENZYME  
 CC USES ERUCOYL-COA AS AN ACYL DONOR.  
 CC -1- CATALYTIC ACTIVITY: Acyl-CoA + 1-acyl-sn-glycerol 3-phosphate -  
 CC CoA + 1,2-diacyl-sn-glycerol 3-phosphate.  
 CC -1- PATHWAY: SECOND STEP IN DE NOVO PHOSPHOLIPID BIOSYNTHESIS.  
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein (Potential).  
 CC -1- SIMILARITY: BELONGS TO THE 1-ACYL-SN-GLYCEROL-3-PHOSPHATE  
 CC ACYLTRANSFERASE FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL: X83266; CAA58239.1; -  
 DR EMBL: Z46836; CAA68877.1; -  
 DR InterPro: IPR002123; Acyltransferase.  
 DR Pfam: PF01553; Acyltransferase; 1.  
 KW Phospholipid biosynthesis; Transferase; Acyltransferase;  
 KW Transmembrane.  
 FT TRANSMEM 40 60 POTENTIAL.  
 FT TRANSMEM 71 91 POTENTIAL.  
 FT TRANSMEM 110 130 POTENTIAL.  
 FT CONFLICT 46 46 I -> V (IN REF. 2).  
 FT CONFLICT 188 188 R -> G (IN REF. 2).  
 FT CONFLICT 262 262 V -> I (IN REF. 2).  
 FT CONFLICT 281 281 N -> K (IN REF. 2).  
 SQ SEQUENCE 281 AA; 31716 MW; 9C880BD9E492EE2A CRC64;

Query Match 6.4%; Score 119; DB 1; Length 281;  
 Best Local Similarity 23.8%; Pred. No. 0.0062;  
 Matches 65; Conservative 40; Mismatches 92; Indels 76; Gaps 14;

QY 2 RYLLPSV-----VLGTAPTYV-----LAWGVRLSLFAPRFQAL 39  
 DB 13 KOLKPVNATADDDKGVNVLSCRFICFAIVLITAAWGL--IMVLLPWPYRIR 70  
 QY 40 DRLRY-CVYOSVLPFFENYTVQVLIYGLDPKKNENIYLANHSTVDVADILAIRQ 98  
 DB 71 LGLYGHIIIGGLVIMY----GIPKIQGS-EHTKRAIYISNHASIDAFVFWMLA--- 122  
 QY 99 NALGHVRYVLKEGLKMLPLYGCTFAOHGITYVRSKAFENKRNKLOSYVDAGTP--MY 156  
 DB 123 -PIGTGVAKKEVI-WYPLLGLQYTLAHNIRIDRS--NPAALIOSKKEAVRITERNKLS 177  
 QY 157 LVYFPGSTRNPEQTVLASOAF---AAQRGLAVLKHVLTPIRIKATHVAF----- 204  
 DB 178 LIMEFPGSTR--SHDGRLLPDKKGFVHLAQSHLPVPMILT-----GTHLAWRKGTFRVRP 231  
 QY 205 -----DCMKNYLDAYDVTV 219  
 DB 232 VPIYVYLPPIINTDWTVDKIDYVAKMHIDVYV 264  
 RESULT 13  
 PLSC\_LIML  
 ID PLSC\_LIML STANDARD; PRT; 281 AA.  
 AC 042868;  
 DT 15-DEC-1998 (Rel. 37, Created)  
 DT 15-DEC-1998 (Rel. 37, Last sequence update)  
 DE 1-acyl-sn-glycerol-3-phosphate acyltransferase (EC 2.3.1.51) (1-AGP  
 DE acyltransferase) (1-AGPAT) (Lysophosphatidic acid acyltransferase)  
 DE (LPAAT).  
 OS Limnantes alba (White meadowfoam).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
 OC eurosids II; Brassicales; Limnathaceae; Limnantes.  
 OX NCBI\_TaxID=42439;  
 RN [1]  
 RC SEQUENCE FROM N.A.  
 RC STRAIN=ALBA;  
 RX MEDLINE=96106201; PubMed=8539298;  
 RA Lasserre M.W., Levering C.K., Davies H.M.D., Knutson D.S.;  
 RT "Lysophosphatidic acid acyltransferase from meadowfoam mediates  
 RT insertion of erucic acid at the sn-2 position of triacylglycerol in  
 RT transgenic rapeseed oil.";  
 RL Plant Physiol. 109:1389-1394(1995).  
 CC -1- FUNCTION: CONVERTS LYOPHOSPHATIDIC ACID (LPA) INTO PHOSPHATIDIC  
 CC ACID BY INCORPORATING ACYL MOIETY AT THE 2 POSITION. THIS ENZYME  
 CC USES ERUCOYL-COA AS AN ACYL DONOR.  
 CC -1- CATALYTIC ACTIVITY: Acyl-CoA + 1-acyl-sn-glycerol 3-phosphate -  
 CC CoA + 1,2-diacyl-sn-glycerol 3-phosphate.  
 CC -1- PATHWAY: SECOND STEP IN DE NOVO PHOSPHOLIPID BIOSYNTHESIS.  
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein (Potential).  
 CC -1- SIMILARITY: BELONGS TO THE 1-ACYL-SN-GLYCEROL-3-PHOSPHATE  
 CC ACYLTRANSFERASE FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL: U32968; AAC49185.1; -  
 DR InterPro: IPR002123; Acyltransferase.  
 DR Pfam: PF01553; Acyltransferase; 1.  
 KW Phospholipid biosynthesis; Transferase; Acyltransferase;  
 KW Transmembrane.  
 FT TRANSMEM 40 60 POTENTIAL.  
 FT TRANSMEM 71 91 POTENTIAL.  
 FT TRANSMEM 110 130 POTENTIAL.  
 SQ SEQUENCE 281 AA; 31664 MW; 2725B811777A6067 CRC64;

Query Match 6.2%; Score 116; DB 1; Length 281;  
 Best Local Similarity 24.2%; Pred. No. 0.011;  
 Matches 59; Conservative 41; Mismatches 86; Indels 58; Gaps 13;

QY 9 VILGTAPTYVLAAGVRLSAPLAPRFQALDRLT-CVYOSVLPFFENYTVQVLIYGL 67  
 DB 46 IYLVTA-----VAMGL--IMVLLPWPYRIRLGNLYGIIIGLVIMY----GIPKIQGS 95  
 QY 68 DLPKKNENIYLANHSTVDVADILAIRONALGHVRYVLKEGLKMLPLYGCTFAOHGG 127  
 DB 96 S-EHTKRAIYISNHASPIDAFVFWMLA---PIGTGVAKKEVI-WYPLLGLQYTLAHN 149  
 QY 128 IYVRSKAFENKRNKLOSYVDAGTP--MYLVIFPGSTRNPEQTVLASOAF---AA 182  
 DB 150 IRLDRS--NPAALIOSKKEAVRITERNKLSLIMEFPGSTRG--DGRLLPDKKGFVHL 204  
 QY 183 ORGLAVLKHVLTPIRIKATHVAF-----DCMKNYLDAYDVTV 215  
 DB 205 QSHLPVPMILT-----GTHLAWRKGTFRVRPVIYKYLPPIINTDWTVDKIDYVAKMH 260  
 QY 216 DVTV 219  
 DB 261 DIVV 264  
 RESULT 14  
 PLSC\_MYCPN  
 ID PLSC\_MYCPN STANDARD; PRT; 266 AA.  
 AC P75479;  
 DT 01-NOV-1997 (Rel. 35, Created)  
 DT 01-NOV-1997 (Rel. 35, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)

Query Match	5.18;	Score 114;	DB 1;	Length 266;
Best Local Similarity	25.58;	Pred. No. 0.016;		
Matches	56;	Conservative	42;	Mismatches 74; Indels 48; Gaps 13;

  

QY	1	MRYL-LPSVLLGTAPTYVLAMGVRLLSAPLPAHFQALD-----RLYCYQSYWLF	53
Db	13	LRFLQLSLVLY--DPVFL-----MLISLISAKNYESIPENPPETIRFKKVVRLVSF	65
QY	54	FFENTGYOILLGDLPRN--KENIITYLANHOSYDWMIVADILAIROMALGHVRYLKEG	111
Db	66	LY--IKGKRVIVN--PENVRKKAVLVVANHKSNDLPILLIKAFKTEGSPPLITIAKIE	121
QY	112	LK--WL-----PLYGCFPAOHGGIYVVKRSAKENEKEMRNKL--QSYVDAGTMYLVIFE	162
Db	122	LQDTWLFKIMKILDFVFD-----RKNLRQMASLSEQOQOIIKRGTA--LCVEPE	169
QY	163	GTRYNPEQTKVLSAQAFQAORGLVAVLKHVLTPIRKATNV	202
Db	170	GTR-----VLSRQIGEFKSGALKVATNAFPIPLTIV	202

  

RESULT	15		
EXON_HSV6U	STANDARD;	PRT;	488 AA.
AC	P24447;		
DT	01-MAR-1992 (Rel. 21, Created)		
DT	01-MAR-1992 (Rel. 21, Last sequence update)		
DT	16-OCT-2001 (Rel. 40, Last annotation update)		
DE	Alkaline exonuclease (EC 3.1.11.-).		
GN	U70 OR 16R.		
OS	Human herpesvirus (type 6 / strain Uganda-1102) (HHV6).		
OC	Vituses; dsDNA viruses, no RNA stage; Herpesviridae;		

[illegible]

(

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 28, 2002, 11:07:36 ; Search time 65.95 Seconds  
(without alignments)  
925.962 Million cell updates/sec

Title: US-09-853-526-4  
Perfect score: 1871  
Sequence: 1 MRYLPSVVLGTAPTYVLA.....YVNTWYGTLLGLMTYKA 353

Scoring table: BIOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 562222 seqs, 172994929 residues

Total number of hits satisfying chosen parameters: 562222

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

- 1: SP-archaea:\*
- 2: SP-bacteria:\*
- 3: SP-fungi:\*
- 4: SP-human:\*
- 5: SP-invertebrate:\*
- 6: SP-mammal:\*
- 7: SP-mhc:\*
- 8: SP-organelle:\*
- 9: SP-phage:\*
- 10: SP-plant:\*
- 11: SP-rodent:\*
- 12: SP-virus:\*
- 13: SP-vertebrate:\*
- 14: SP-unclassified:\*
- 15: SP-virus:\*
- 16: SP-bacteriap:\*
- 17: SP-archaeap:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	% Match	Query Length	DB ID	Description
1	1871	100.0	364	4 Q9BG04	Q9bg04 homo sapien
2	1868	99.8	353	4 Q9NUQ2	Q9nuq2 homo sapien
3	1568.5	83.8	354	11 Q9D1B8	Q9d1b8 mus musculu
4	453	24.2	386	5 Q9SRI2	Q9sri2 caenorhabdi
5	398.5	21.3	373	10 Q9SSHO	Q9ssho arabidopsis
6	398.5	21.3	393	10 Q9C9P8	Q9c9p8 arabidopsis
7	390.5	20.9	375	10 Q9LHNA	Q9lhna arabidopsis
8	328	17.5	374	10 Q4I745	Q4i745 zea mays (m
9	309.5	16.5	390	10 Q9XFW4	Q9xfw4 brassica na
10	307	16.3	377	10 Q40119	Q40119 limnanthes
11	305.5	16.3	306	10 Q9SDN3	Q9sdn3 prunus dulc
12	272.5	14.6	386	5 Q9VY51	Q9vy51 drosophila
13	270	14.4	378	11 Q9Z4S1	Q9z4s1 rattus norv
14	267	14.3	380	5 Q9VY49	Q9vy49 drosophila
15	263	14.1	376	10 Q9SYC8	Q9sy8 arabidopsis
16	251	13.4	316	11 Q9DB84	Q9db84 mus musculu

17	249.5	13.3	350	3 Q94361	Q94361 schizosacch
18	242.5	13.0	344	5 Q16526	Q16526 caenorhabdi
19	239	12.8	311	10 Q93317	Q93317 brassica na
20	238.5	12.7	310	10 Q9SVX9	Q9svx9 arabidopsis
21	223	11.9	314	11 Q9D517	Q9d517 mus musculu
22	214	11.4	399	5 Q23087	Q23087 caenorhabdi
23	212	11.3	439	5 Q20800	Q20800 caenorhabdi
24	205.5	11.0	370	11 Q91YX5	Q91yx5 mus musculu
25	183.5	9.8	426	5 Q01882	Q01882 caenorhabdi
26	183.5	9.8	426	5 Q9SR03	Q9sr03 caenorhabdi
27	178	9.5	295	16 Q91657	Q91657 pseudomonas
28	152.5	8.2	304	16 Q91546	Q91546 pseudomonas
29	145	7.7	359	10 Q9PF57	Q9pf57 arabidopsis
30	131.5	7.0	363	5 Q19221	Q19221 caenorhabdi
31	130	6.9	294	16 Q9KVV5	Q9kvv5 vibrio chol
32	124	6.6	248	2 Q30849	Q30849 salmonella
33	116	6.2	281	10 Q9SDQ2	Q9sdq2 limnanthes
34	114.5	6.1	247	16 Q9X219	Q9x219 thermotoga
35	113.5	6.1	385	5 Q9NFP1	Q9nfp1 leishmania
36	113	6.0	262	16 Q98G17	Q98g17 rhizobium 1
37	112	6.0	241	16 Q9CPE2	Q9cpe2 pasteurella
38	112	6.0	258	16 Q9HW50	Q9hw50 pseudomonas
39	109.5	5.9	235	2 Q9KX14	Q9kx14 rhodobacter
40	109.5	5.9	488	12 Q9WT03	Q9wt03 human herpes
41	108	5.8	261	16 Q9KPE3	Q9kpe3 vibrio chol
42	105.5	5.6	249	2 Q9EY25	Q9ey25 xanthomonas
43	105	5.6	260	16 Q9ZMG4	Q9zmg4 rhizobium m
44	104	5.6	344	10 Q9LLY4	Q9lly4 brassica na
45	102.5	5.5	291	3 Q9Y7C5	Q9y7c5 emeritella

## ALIGNMENTS

RESULT	ID	PRELIMINARY:	PRT:	364 AA.
Q9BG04	Q9BG04	Q9BG04		
AC	Q9BG04	Q9BG04		
DT	01-JUN-2001 (TREMBLrel. 17, Created)			
DT	01-JUN-2001 (TREMBLrel. 17, Last sequence update)			
DT	01-DEC-2001 (TREMBLrel. 19, Last annotation update)			
DE	HYPOTHETICAL 42.1 KDA PROTEIN (LYSOPHOSPHATIDIC ACID			
DE	ACYLTRANSFERASE-EPSILON) (EC 2.3.1.51).			
GN	DKFP761C222 OR LPAAT-E.			
OS	Homo sapiens (Human)			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
OX	NCBI_TaxID=9606;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=AMYGDALA;			
RX	MEDLINE=21154917; PubMed=11230166;			
RA	Wiemann S., Weill B., Wellenreuther R., Gassenhuber J., Glass S.,			
RA	Ansorge W., Boecker M., Bloecher H., Bauersachs S., Blum H.,			
RA	Lauber J., Duesterhoeft A., Beyer A., Koehrer K., Struck N.,			
RA	Mewes H.W., Oltersdoerfer B., Obermaier B., Tampe J., Heubner D.,			
RA	Mambutt R., Korn B., Klein M., Poustka A.,			
RT	"Towards a Catalog of Human Genes and Proteins: Sequencing and			
RT	Analysis of 500 Novel Complete Protein Coding Human cDNAs.";			
RL	Genome Res. 11:422-435(2001).			
RN	[2]			
RP	SEQUENCE FROM N.A.			
RA	Leung D.W.;			
RT	"Cloning and expression of LPAAT-epsilon.";			
RT	Submitted (May-2001) to the EMBL/Genbank/DBJ databases.			
DR	EMBL; AL136587; CAB66522.1; -			
DR	EMBL; AF375789; AAK54809.1; -			
DR	InterPro: IPR002123; Acyltransferase.			
DR	Pfam: PF01553; Acyltransferase; 1.			
KW	Hypothetical protein; transferase; Acyltransferase.			
SQ	SEQUENCE 364 AA; 42072 MW; 90A0F87FC7C8081 CRC64;			

Query Match	100.0%	Score 1871;	DB 4;	Length 364;
Best Local Similarity	100.0%;	Pred. No. 4,2e-166;		
Matches 353;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
QY	1	MRLLPSVLLGCTAFTYVLAMGVMRLLSAFLRAREFOALDRLRLCYGQSMVLEFFENYTG	60	
Db	12	MRLLPSVLLGCTAFTYVLAMGVMRLLSAFLRAREFOALDRLRLCYGQSMVLEFFENYTG	71	
QY	61	VOILLYGDLPRKKEKITIYLANHQSTVDIVADILAIRONALGHVRRVYLKEGKLWLPYLCG	120	
Db	72	VOILLYGDLPRKKEKITIYLANHQSTVDIVADILAIRONALGHVRRVYLKEGKLWLPYLCG	131	
QY	121	YFAOHGGIYVKSANAFNEKEEMKNKILQSYVDATPMATVLYFERGSTRNPNQOTVLSAQAF	180	
Db	132	YFAOHGGIYVKSANAFNEKEEMKNKILQSYVDATPMATVLYFERGSTRNPNQOTVLSAQAF	191	
QY	181	AAORGLAVLAKHVLTPRIKATHVAFLFCMKNYLDAIVDYVVEEGKDGCGRRSPPTTEFL	240	
Db	192	AAORGLAVLAKHVLTPRIKATHVAFLFCMKNYLDAIVDYVVEEGKDGCGRRSPPTTEFL	251	
QY	241	CKECPKIHIIHIDRIDKKDVPPEQEHMRRLVHERFEIKDKMLIEFYESPDERRRKRPCKGS	300	
Db	252	CKECPKIHIIHIDRIDKKDVPPEQEHMRRLVHERFEIKDKMLIEFYESPDERRRKRPCKGS	311	
QY	301	VNSKLSIKKTLPSMLISGLTAGMLMTAGAKRLVYNTWTYIGLLGCLMWTTIKA	353	
Db	312	VNSKLSIKKTLPSMLISGLTAGMLMTAGAKRLVYNTWTYIGLLGCLMWTTIKA	364	

RESULT	2			
09NM02				
ID	Q9NM02	PRELIMINARY;	PRF:	353 AA.
AC	Q9NM02;			
DT	01-OCT-2000	(TREMBLrel. 15, Created)		
DT	01-OCT-2000	(TREMBLrel. 15, Last sequence update)		
DT	01-JUN-2001	(TREMBLrel. 17, Last annotation update)		
DE	CDNA FLJ11210 F1S, CLONE PLACE1007954.			
OS	Homio sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.			
OX	NCBI_TaxID=9606;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=PLACENTA;			
RA	Isoegal T., Oca T., Hayashi K., Sugiyama T., Otsuki T., Suzuki Y.,			
RA	Nishikawa T., Nagai K., Sugano S., Shiratori A., Sudo H.,			
RA	Wagatsuma M., Hosoki T., Kaku Y., Kodaira H., Kondo H., Sugawara M.,			
RA	Watanabe S., Chiba Y., Ishida S., Murakawa K., Ono Y., Takiguchi S.,			
RA	Yamanoto J., Wakematsu A., Nakamura Y., Nagahari K., Masuno Y.,			
RA	Ninomiya K., Iwayanagi T.;			
RT	"NEDO human cDNA sequencing project.";			
RL	Submitted (Feb-2000) to the EMBL/Genbank/DBJ databases.			
DR	EMBL; AK002072; BAA92069.1;			
DR	InterPro: IPR002123; Acyltransferase.			
DR	Pfam: PF01553; Acyltransferase; 1.			
SQ	SEQUENCE 353 AA; 40799 MW; 0A8A14A2361FB14 CRC64;			

[illegible]

Db	121	YFAGGAGIYVRSAAFNENKEMRNKQSVVDAGTGPRLYLIFIEPGTRINPEQRIKVLASAGAF	180
Oy	181	AAQGLAVLKHVLPRIKATHVAFDCKMKNYDAIYDVTVYEGKDDGQRESPTMTFFL	240
Db	181	AAQGLAVLKHVLPRIKATHVAFDCKMKNYDAIYDVTVYEGKDDGQRESPTMTFFL	240
Oy	241	CKECKIHIIHDIRDKVPEBOEHMRWMLHERPEIKKMLIEFYESPDPERRRRFPCKS	300
Db	241	CKECKIHIIHDIRDKVPEBOEHMRWMLHERPEIKKMLIEFYESPDPERRRRFPCKS	300
Oy	301	VNSKLSIKKTLPSMLILSGLTAGMLMTBAGRKLYNWTIYGTLLGCLMTVITKA	353
Db	301	VNSKLSIKKTLPSMLILSGLTAGMLMTBAGRKLYNWTIYGTLLGCLMTVITKA	353
RESULT	3		
O9DIE8			
ID	O9DIE8	PRELIMINARY;	PRT; 354 AA.
AC	O9DIE8:		
DT	01-JUN-2001 (TREMBLrel. 17, Created)		
DT	01-JUN-2001 (TREMBLrel. 17, Last sequence update)		
DT	01-JUN-2001 (TREMBLrel. 17, Last annotation update)		
DE	1110013A05RIK PROTEIN.		
GN	1110013A05RIK.		
	Mus musculus (Mouse).		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.		
NCBI	taxid=10090;		
NN	[1]		

RP SEQUENCE FROM N.A.: TISSUE=EMBRYO;  
RC STRAIN=C57BL/6J; MEDLINE=21085660; Pubmed=11217851;  
RX MEDLINE=21085660; Pubmed=11217851;  
RA Kawai J., Shinaigawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,  
RA Arikawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,  
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamamaka I.,  
RA Saito T., Okazaki Y., Gojobori T., Bono H., Ksukawa T., Saito R.,  
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,  
RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochwa H.,  
RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,  
RA Schirral L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,  
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,  
RA Blake J., Botfell D., Bojunga N., Carlinici P., de Bonaldo M.F.,  
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,  
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,  
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Momhaerts P.,  
RA Nardone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,  
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,  
RA Suzuki H., Toyooka K., Wang K.H., Welts C., Whitaker C., Wilmink L.,  
RA Wyszynski-Boris A., Yoshida K., Hasegawa Y., Kawai H., Kohsaki S.,  
RA Hayashizaki Y.;  
RT "Functional annotation of a full-length mouse cDNA collection.";  
RL Nature 409:685-690(2001).  
DR EMBL: AK003649; BAB22915.1; -;  
DR MGD: MGI:1915880; 1110013A05RLK.  
DR InterPro: IPR002123; Acyltransferase.  
pfam: PF01553; Acyltransferase; 1.  
QO SEQUENCE 354 AA; 40943 MW; 447EEF924B91E800 CRC64;

[illegible]



```

Db 121 YFAHGIGIYVKSAAKFNDEKMSKLSQSYVNAGMTMYLVEFPEGTRYNATYTKLISASQAF 180
Qy 161 AAORGLAVLKHLVTPRIKATHAFFDCMKNYLDAIDYDVVVEGKDDG-CQRESTWTFEF 239
Db 161 AAORGLAVLKHLVTPRIKATHAFFDCMKSHLDAYDVTVVEGNGKSGKTSNPSMTFEEF 240
Qy 240 LKEGCPKIHIDRIDKDVPEDEQEHMRMWLHEREFEIKDKMLIEFESPDDEPRRRFRFGK 299
Db 241 LCKGCPKLIHFDRIDRNEVEPEQEHMRMKWLHEREFEIKDRILLIETDSDPDERRRKFFRGK 300
Qy 300 SVNSKLSIKTKLPSPMLISGLTAGMLMTDAGRKLYVNTWIXYTGTLGCLMTVITKA 353
Db 301 SVHSRLSVKTLPSVLLIGSLTAVVLMTESGKLYMGTWGLTGCLMFWVKA 354

```

```

RESULT 4
ID Q95R12 PRELIMINARY; PRT; 386 AA.
AC Q95R12;
DT 01-DEC-2001 (TREMBLrel. 19, Created)
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE HYPOTHEICAL 44.0 KDA PROTEIN.
GN F28B3.9.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-BRISTOL N2;
RX MEDLINE=99069613; PubMed=9851916;
RA None;
RT "Genome sequence of the nematode C. elegans: a platform for
RT investigating biology. The C. elegans Sequencing Consortium.";
RL Science 282:2012-2018(1998).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-BRISTOL N2;
RA Gelsel C., Kramer J., Smith A.;
RT "The sequence of C. elegans cosmid F28B3.";
RT Submitted (May-1997) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN-BRISTOL N2;
RA Waterston R.;
RT "Direct Submission.";
RL EMBL: AF003136; AK93853.1; -
KW Hypothetical protein.
SQ SEQUENCE 386 AA; 43978 MW; A1E3020C201452A6 CRC64;

```

Query Match 24.2%; Score 453; DB 5; Length 386;  
 Best Local Similarity 32.9%; Pred. No. 8.1e-34;  
 Matches 127; Conservative 72; Mismatches 139; Indels 48; Gaps 14;

```

Qy 1 MYLLP-----SVLLGTAPTYVLANGVWRLLSAFLPARFYQALDRRLCYVOSMTLFFEE 56
Db 12 LRPILPCSLSLSMVPFASCAIVIGV-----SWIVPRHVAOQDLMMLKYSYWRLLCLFFEE 67
Qy 57 NTGVOILLYGLDLPK--NK-----ENITYLANHSTVDIVADIIAIRONALGH---VRY 107
Db 68 NISGVEITLHGHTNEEVYVKNKTGFENAVMISNHSNVDMIIIPMLARHGDCGNEQAFRYM 127
Qy 108 LKEGLKMLPLGYCYFAOHGIGIYVKSAAKFNDEKMSKLSQSYVNAGMTMYLVEFPEGTRYN 167
Db 128 VKNSTIHLVPMFGWYLFQGYIVRRFEPGAPVLRLQKWLNESDPYWLILFFEGTRNS 187
Qy 168 PEQTVLASQAFQAORGLAVLKHLVTPRIKATHAFFDCMKNYLDAIDYDVVY-EGKDD 226
Db 188 AKKHLLSSNRFLEKSGROPQWNVLCPRSGLALALDNLST-LDAIDYDVTVYGMQMRD 246

```

```

Qy 227 GG-----QRESPTWTFELC--KECPKIHIDRIDKDVPEDEQEHMR 267
Db 247 LGLNLTINYNLIIYRMAERKGLAPGMDPCCGQFQGLIHIDRIDIDEVPRAKLELR 306
Qy 268 RWLHEREFIKDKMLIEFESPDDEPRRRFRFGKSVNSKLSIKTKLPSPMLIS-GTAGMLM 326
Db 307 TWTLERFTRKRIIDEFY-SEKPSGSAIP-----CVPISQTLPSLTFESALLAPFFS 359
Qy 327 TDAGRKLYVNTWIXYTGTLGCLMTVITK 352
Db 360 RTIGR-IYLLTASSPL-IAMLHIR 383

```

```

RESULT 5
ID Q9SSH0 PRELIMINARY; PRT; 373 AA.
AC Q9SSH0;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
DE F25A4.2 PROTEIN.
GN F25A4.2.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC Eucosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-CV. COLUMBIA;
RA Vysotskaya V.S., Schwartz J.R., Yu G., Toriumi M., Lenz C., Liu S.,
RA Lee J.M., Li J., Gonzalez A., Liu A., Liu K., Vaysberg M., Sakano H.,
RA Chin C., Choi E., Chlou J., Altati H., Araujo R., Brooks S.,
RA Buehler E., Chao Q., Conn L., Conway A.B., Dunn P., Hansen N.,
RA Hwang B., Huizlar L., Khan S., Kim C., Palm C., Rowley D., Shinn P.,
RA Walker M., Davis R.W., Ecker J.R., Federspiel N.A., Theologis A.;
RT "Arabidopsis thaliana chromosome 1 BAC F25A4 sequence.";
RL Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL: AC008263; AAD55275.1; -
DR InterPro: IPR002123; Acyltransferase.
DR Pfam: PF01553; Acyltransferase; 1.
SQ SEQUENCE 373 AA; 42383 MW; 98642948D94709B6 CRC64;

```

Query Match 21.3%; Score 398.5; DB 10; Length 373;  
 Best Local Similarity 29.8%; Pred. No. 9.4e-29;  
 Matches 97; Conservative 68; Mismatches 117; Indels 43; Gaps 7;

```

Qy 1 MYLLPVSLLGTAPTYVLANGVWRLLSAFLPARFYQALDRRLCYVOSMTL----- 52
Db 23 LRGLMILLVFLSTAFMFL-----YFAP---IALGLRLLSVQSKRYVSLIFGLM 70
Qy 53 -----FFENYTGVOILLYGLDLPKKNENIYLANHSTVDIVADIIAIRONALGHVRY 107
Db 71 LALMPLRETVANGTVVFSGDIIPEKRVLLIANHRTVDWMYIMNIALRKGCGLYRYV 130
Qy 108 LKEGLKMLPLGYCYFAOHGIGIYVKSAAKFNDEKMSKLSQSYVNAGMTMYLVEFPEGTRYN 167
Db 131 LKSSLMKLPFRFGWGHVLEFIPVERKRRVDEVLLQMLSSFDPOEPLMLLFFEGTFT 190
Qy 168 PEQTVLASQAFQAORGLAVLKHLVTPRIKATHAFFDCMKNYLDAIDYDVVVEGKDDG 227
Db 191 EKKCK---RSQFAAEVGLPALSNVLLPKTRGFCLEVLNNSLDAYVDLITAKPR--- 244
Qy 228 GQRESPTWTFELKECP-KIHIDRIDKDVPEDEQEHMRMWLHEREFEIKDKMLIEYE 286
Db 245 -----CPSFMDVVFEGTDSSEVHIHVRVLKEIPANEAESSAMLDSPFKLDKLSDF-- 297
Qy 267 SPDERKRRFRFGKSVNSKLSIKTKL 311
Db 298 ---NAQGFPMQOREEELSVLKCI 318

```

RESULT 6  
Q9C9P8 PRELIMINARY: PRT: 393 AA.  
AC Q9C9P8: 01-JUN-2001 (TREMBLrel. 17, Created)  
DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)  
DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)  
DE PUTATIVE ACYL-COA:1-ACYLGLYCEROL-3-PHOSPHATE ACYLTRANSFERASE, 31588-29381.  
GN F9E10.13.  
OS Arabidopsis thaliana (Mouse-ear cress).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
OC eucosids II; Brassicales; Brassicaceae; Arabidopsis.  
OX NCBI\_TaxID=3702;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-CV. COLUMBIA;  
RA MEDLINE=21016719; PubMed=11130712;  
RA Theologis A., Ecker J.R., Palm C.J., Federspiel N.A., Kaul S.,  
RA White O., Alonso J., Altafi H., Araujo R., Bowman C.L., Brooks S.Y.,  
RA Buehler E., Chan A., Chao Q., Chen H., Cheuk R.F., Chin C.W.,  
RA Chung M.K., Conn L., Conway A.B., Cressy T.H., Dewar K.,  
RA Dunn P., Eguu P., Feldblyum T.V., Feng J.-D., Fong B., Fujii C.Y.,  
RA Gill J.E., Goldsmith A.D., Haas B., Hansen N.F., Hughes B., Huizar L.,  
RA Hunter J.L., Jenkins J., Johnson-Hopson C., Kurtz D.B., Kwan A., Lam B.,  
RA Kim C.J., Koo H.L., Kremetska I., Kurtz D.B., Kwan A., Lam B.,  
RA Langin-Hooper S., Lee A., Lee J.M., Lenz C.A., Li J.H., Li Y.-P.,  
RA Lin X., Liu S.X., Liu Z.A., Luros J.S., Maiti R., Marzilli A.,  
RA Miltsecher J., Miranda M., Nguyen M., Nierman W.C., Osborne B.I.,  
RA Pal G., Peterson J., Pham P.K., Rizzo M., Rooney T., Rowley D.,  
RA Sakano H., Salzberg S.L., Schwartz J.R., Shinn P., Southwick A.M.,  
RA Sun H., Tallon L.J., Tambunga G., Toriumi M.J., Town C.D.,  
RA Uterback T., Van Aken S., Vaysberg M., Vysotskaia V.S., Walker M.,  
RA Wu D., Yu G., Fraser C.M., Ventner J.C., Davis R.W.;  
RT "Sequence and analysis of chromosome 1 of the plant Arabidopsis  
thaliana";  
RL Nature 408:816-820(2000).  
DR EMBL: AC013258; AAG51931.1;  
DR InterPro: IPR002123; Acyltransferase.  
DR Pfam: PF01553; Acyltransferase; 1.  
KW Transferase; Acyltransferase.  
SQ SEQUENCE 393 AA; 44892 MW; FA5949D21E8E9A29 CRC64;

Query Match 21.3%; Score 398.5; DB 10; Length 393;  
Best Local Similarity 29.8%; Pred. No. 1e-28;  
Matches 97; Conservative 68; Mismatches 117; Indels 43; Gaps 7;  
QY 1 MYRLPSVLLGTAFTYVLAAGVWLLSAFLPARFYQALDDRLCYVQSNVLF----- 52  
DB 23 LRGLMILLVFLSTAFMFL-----YFAP--IAALGLRLSVQGRKVVSLIFGLM 70  
QY 53 -----FFFNYYGVOILLXGDLPRKNENIYLANHSTVDWIVADILAIKONALGHVRY 107  
DB 71 LALMPILFETVNGTIVFSGDIIIVKRVLLANHRTEDMMLMIALRKGLGKIVKY 130  
QY 108 LKEGLKMLPLGCGYFAOHGIIYKRSKFNKEMRNKLOSVDAGPMTLVIPPESTRN 167  
DB 131 LKSSLMKLLIFGNGFVLEFIPERKREVDKLVLOMLSSFKDPOEPLMALPPEGTDT 190  
QY 168 PEOTKVLSQAFAAGRLAVLKHVLPRIKATHVAFDCKNVLDAIYDVTVVECKDGG 227  
DB 191 EERCK---RSOKFAAEVGLPALSNNVLPKTRGCGVCLVLEHNSLDLVYDTIAVYKR--- 244  
QY 228 GORRESPTWTEFLCKECP-KIHIIHDIRDKDQVPEQEHMRRLHREFEIKKMLIE 286  
DB 245 -----CPSPMDNVYFGTDPSEVHIVHVRVLIKEIPANEAESSAAMLSDFKDKLSDF-- 297  
QY 287 SPDPERKRRPFGKSVSKLSIKKTL 311  
DB 298 ----NAQGRFPNORPEELSLVKCI 318

RESULT 7  
Q9LHN4 PRELIMINARY: PRT: 375 AA.  
AC Q9LHN4: 01-OCT-2000 (TREMBLrel. 15, Created)  
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)  
DT 01-DEC-2001 (TREMBLrel. 15, Last annotation update)  
DE ACYL-COA:1-ACYLGLYCEROL-3-PHOSPHATE ACYLTRANSFERASE.  
OS Arabidopsis thaliana (Mouse-ear cress).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
OC eucosids II; Brassicales; Brassicaceae; Arabidopsis.  
OX NCBI\_TaxID=3702;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-COLUMBIA;  
RA Kaneko T., Kato T., Sato S., Nakamura Y., Asamizu E., Tabata S.;  
RL Submitted (MAY-2000) to the EMBL/Genbank/DBJ databases.  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN-COLUMBIA;  
RX MEDLINE=20363099; PubMed=10907853;  
RA Nakamura Y.;  
RT "Structural analysis of Arabidopsis thaliana chromosome 3. II.  
RT Sequence features of the regions of 4,251,695 bp covered by ninety pl,  
RT TAC and BAC clones";  
RL DNA Res. 7:217-221(2000).  
DR EMBL: AP002039; BAB03094.1;  
DR InterPro: IPR002123; Acyltransferase.  
DR Pfam: PF01553; Acyltransferase; 1.  
KW Transferase; Acyltransferase.  
SQ SEQUENCE 375 AA; 43490 MW; 8F6F5DC313A3D704 CRC64;

Query Match 20.9%; Score 390.5; DB 10; Length 375;  
Best Local Similarity 26.9%; Pred. No. 5.3e-28;  
Matches 101; Conservative 78; Mismatches 128; Indels 69; Gaps 9;  
QY 1 MYRLPSVLLGTAFTYVLAAGVWLLSAFLPARFYQALDDRLCYVQSNVLF----- 55  
DB 18 LRGLMILLVFLSTAFMFLFNGFSLAVVLRFS-----IRYSKRVSPFGSLALMPFLF 73  
QY 56 ENYTGVOILLYDLPKKNENIYLANHSTVDWIVADILAIKONALGHVRYLAKGLKWL 115  
DB 74 EKINKKRVIFSGDKVCEEDRVLLINHRTEDMMLMIALRKGLGKIVKY 133  
QY 116 PLYGCYFAOHGIIYKRSKFNKEMRNKLOSVDAGPMTLVIPPESTRNPEQTKVLS 175  
DB 134 PLEGAHFILFEFTPVKRWVEDEANLROIVSSFKDPRDALMLALPPEGDIYEAQC--- 190  
QY 176 ASQAFAGRLAVLKHVLPRIKATHVAFDCKNVLDAIYDVTVVECKDGGORRESPT 235  
DB 191 RSKKFAENGDLILNNVLPKTRGCVSCQELSCGLDANVYDTIYKTR----- 239  
QY 236 MTEFLCKECP-----KIHIIHDIRDKDQVPEQEHMRRLHREFEIKKMLIE 283  
DB 240 -----CPSLDNVYGIPESEVHIIHRIINLITOIPNOEKDINAMLNTFOLKDLIND 291  
QY 284 FYESPDPERKRRPFGKSVSKLSIKKTLPSMLISGLT-----AGMLTMDGRRLY 334  
DB 292 FYSN-----GHFPNEGKEKENTRKYLLINCLAVIAFTTICHTLFFSSMWF-----RIY 341  
QY 335 VMTWITIGTLGLAWT 350  
DB 342 VS-----LACVYLT 350  
RESULT 8  
Q41745 PRELIMINARY: PRT: 374 AA.  
AC Q41745: 01-NOV-1996 (TREMBLrel. 01, Created)  
DT

DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)  
 DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)  
 DE 1-ACYL-GLYCEROL-3-PHOSPHATE ACYLTRANSFERASE (PUTATIVE).  
 OS Zea mays (Maize).  
 CC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 CC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC clade;  
 CC Panicoideae; Andropogoneae; Zea.  
 OX NCBI\_TaxID=4577;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=BLACK MEXICAN SWEET; TISSUE=ENDOSPERM;  
 RX MEDLINE=95035993; PubMed=7948871;  
 RA Brown A.P., Coleman J., Tommey A.M., Watson M.D., Slabas A.R.;  
 RT "Isolation and characterization of a maize cDNA that complements a 1-  
 RT acyl sn-glycerol-3-phosphate acyltransferase mutant of E.coli and  
 RT encodes a protein which has similarities to other acyltransferases.";  
 RL Plant Mol. Biol. 26:211-223(1994).  
 DR EMBL: Z29518; CAAB2638.1; -;  
 DR InterPro: IPR002123; Acyltransferase.  
 DR InterPro: IPR003015; HLH\_MYC.  
 DR Pfam: PF01553; Acyltransferase; 1.  
 DR PROSITE: PS00038; HELIX\_LOOP\_HELIX; UNKNOWN\_1.  
 KW Transferase; Acyltransferase.  
 SQ SEQUENCE 374 AA; 42571 MW; F1F5492CAF24F93 CRC64;

Query Match 17.5%; Score 328; DB 10; Length 374;  
 Best Local Similarity 29.8%; Pred. No. 3.5e-22;  
 Matches 92; Conservative 64; Mismatches 121; Indels 32; Gaps 8;

QY 18 VLAMGVWMLLA-----FLPAR-----FYQALDRLCYQSNVLPFFENYTGVO 62  
 DB 9 VLPGLFLFLSLGLYVNAIOAVLFVTIRPFSKFYRINFLAELMLQLVWVDMAGAK 68  
 QY 63 ILIYD-----LPKKNENIILYANHOSTVDWIVADILAIRONALGHVRYLAKGLKWLPLY 118  
 DB 69 VOLHADEETRYSMGKEHLIISNHRSDIDMLIGMTIAQSGCLSTLAWMKSSKFLPYI 128  
 QY 119 G--CYFAOHGGIYVRSKAFNEKEMRNKLOSVDAGTPMYLIFPEGPYNEQIKVISA 176  
 DB 129 GMSMFAEY--LFLERSNAKDEKTLKWLQRLKDPFRPFWLALFVGGREFTPAK--LTA 183  
 QY 177 SOAFNAORGAVLKHLVLPRIKATHVAFDCMKNYLDATVYVVEGKDDGGGRRSEPTM 236  
 DB 184 AQEYASQGLPAPRVNVLPRTKGFVSAYSIMRDFVPAITDTVIYP-KDS-----PQPTM 237  
 QY 237 TEFCKECPKTHIHIDRIDKDVPEQEHMRWLHERFEIKDKMLIEFYESPDPERRKE 296  
 DB 238 LRLKAGSSVHIVRMKRRAMSEMPKSDSDVSKWCDIFAKDALDKHLATGTDFDEIRP 297  
 QY 297 PGKSVNSKL 305  
 DB 298 IGRPVKSL 306

RESULT 9  
 O9XFW4 PRELIMINARY; PRT; 390 AA.  
 AC O9XFW4;  
 DT 01-NOV-1999 (TREMBlrel. 12, Created)  
 DT 01-NOV-1999 (TREMBlrel. 12, Last sequence update)  
 DE 01-DEC-2001 (TREMBlrel. 19, Last annotation update)  
 DE ACYL-COA:1-ACYLGLYCEROL-3-PHOSPHATE ACYLTRANSFERASE (EC 2.3.1.51).  
 GN LPAAT.  
 OS Brassica napus (Rape).  
 CC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 CC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
 CC eurosids II; Brassicales; Brassicaceae; Brassica.  
 OX NCBI\_TaxID=3708;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=DEVELOPING SILIQUES;

RA Graefin zu Muenster A., Wolter F.P., Frenzen M.;  
 RT "A cDNA encoding a microsomal 1-acylglycerol-3-phosphate  
 RT acyltransferase of Brassica napus L.";  
 RL Submitted (MAY-1997) to the EMBL/Genbank/DBJ databases.  
 DR EMBL: Z95637; CAB09138.1; -;  
 DR InterPro: IPR002123; Acyltransferase.  
 DR InterPro: IPR003015; HLH\_MYC.  
 DR Pfam: PF01553; Acyltransferase; 1.  
 DR PROSITE: PS00038; HELIX\_LOOP\_HELIX; UNKNOWN\_1.  
 KW Transferase; Acyltransferase.  
 SQ SEQUENCE 390 AA; 43771 MW; F1446EB30009C37 CRC64;

Query Match 16.5%; Score 309.5; DB 10; Length 390;  
 Best Local Similarity 26.5%; Pred. No. 2e-20;  
 Matches 83; Conservative 75; Mismatches 118; Indels 37; Gaps 9;

QY 58 YTGVOILYGD---LPKKNENIILYANHOSTVDWIVADILAIRONALGHVRYLAKGLK 113  
 DB 63 WAGVKIYVADDEFTRNRMCKEHALVVCNHRSDIDMLGVMTIAQSGCLGSALAVMKSSK 122  
 QY 114 WPLPG--CYFAOHGGIYVRSKAFNEKEMRNKLOSVDAGTPMYLIFPESTRNPEOT 171  
 DB 123 FLPVIGMSMFESEY--LFLERNMARDSTLQGLRLNDFPRFWLALFVEGTRFEAK- 179  
 QY 172 KYLSASOAFNAORGAVLKHLVLPRIKATHVAFDCMKNYLDATVYVVEGKDDGGORR 231  
 DB 180 --LKAQETAAASELPVAPRVNVLPRTKGFVSAYSIMRDFVPAITDTVIYP-----KTS 231  
 QY 232 ESPTMTEFLCKECPKTHIHIDRIDKDVPEQEHMRWLHERFEIKDKMLIEFYESPDPPE 291  
 DB 232 PPTMLRLFKGGPSVHVHAIKCHSKMDLPEPDEIAQWCRDGFVAKDALDKHIANA---- 287  
 QY 292 RKRREPCKSVNSKLSIKTLPSMLISG---LTAGML-----MTDAGRLYNTWYIY 340  
 DB 288 --DTEPGQ---KEONIGRPKISLAVVSWACLLTIGAMFLHMSNLFSSWKIALSAFGL 342  
 QY 341 GTLGCIMWTYKA 353  
 DB 343 GIITTCMOTILNS 355

RESULT 10  
 O40119 PRELIMINARY; PRT; 377 AA.  
 ID O40119;  
 AC O40119;  
 DT 01-NOV-1996 (TREMBlrel. 01, Created)  
 DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)  
 DE 01-DEC-2001 (TREMBlrel. 19, Last annotation update)  
 DE 1-ACYL-SN-GLYCEROL-3-PHOSPHATE ACYLTRANSFERASE (PUTATIVE).  
 OS Limnathes douglasii (Douglas's meadowfoam).  
 CC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 CC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
 CC eurosids II; Brassicales; Limnathaceae; Limnathes.  
 OX NCBI\_TaxID=28973;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC MEDLINE=96046746; PubMed=7579178;  
 RA Brown A.P., Brough C.L., Kroon J., Slabas A.R.;  
 RT "Identification of a cDNA that encodes a 1-acyl-sn-glycerol-3-  
 RT phosphate acyltransferase from limnathes douglasii.";  
 RL Plant Mol. Biol. 29:267-278(1995).  
 DR EMBL: Z48730; CAAB8620.1; -;  
 DR InterPro: IPR002123; Acyltransferase.  
 DR InterPro: IPR003015; HLH\_MYC.  
 DR Pfam: PF01553; Acyltransferase; 1.  
 DR PROSITE: PS00038; HELIX\_LOOP\_HELIX; UNKNOWN\_1.  
 KW Transferase; Acyltransferase.  
 SQ SEQUENCE 377 AA; 42780 MW; 0DECDFE2D39687C CRC64;

Query Match 16.4%; Score 307; DB 10; Length 377;  
 Best Local Similarity 25.5%; Pred. No. 3.3e-20;

```

Matches 91: Conservative 85; Mismatches 131; Indels 50; Gaps 13;

QY 26 LLSAPLPAFY---QALDRLCYVOSMY-----LFEFFNYTGQILLYGDLPR--- 71
   | : | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 21 LVVNFIOAVFYLVYVRISDTRIRNTLVAELMLLELVIMAGVAKVQVLTDTESFRL 80
   | : | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 72 -NKENIYLANHOSYDVIVADILAIRQNALGHVRYLKEGKLWPLYG--CYFAOHGSI 128
   | : | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 81 MCKEHALHLCNHRSDIDMLIGWLAORCCGLSSISVAMKRSKFEPLVICSMWSEY--L 138
   | : | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 129 YKRSKAKFEKEMRNKLSQSYVAGCPMYLVIFESGRYRNPEDQKV--LSASQAFNAQRGLA 187
   | : | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 139 FLERNNAKEMNTLSCGLQNLDPKPEFWLALFEVETRF---TKAKLLAQQEYAAAGLAP 194
   | : | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 188 VLKHLVLPRIKATHVAFDCMKNYLAIYDVTVYEGSKDGGORRESPTTEFLCKECPKI 247
   | : | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 195 VPRNVLPRTKGFVSASVMSRSPVAIYDLTAIP-----KTTEPTMLRLFRGSSSV 248
   | : | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 248 HHHIDRIDKVDPEDEHMRMLHERFEIKDKMLEFYSPDPERKRRPPGKSVNSKLSI 307
   | : | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 249 HVHLKRLMKDLPRKTDGVAQWCKQDFISKD-ALLDKHVAED---TFEGLEVQ---DI 299
   | : | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 308 KKTLSMILSLG---LTAGML-----MTDAGRKLVTWMTYGLGCIWNTIKA 353
   | : | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 300 GRPKMSLVVAVSMKCLLGLVKKFLQWSALSSWKGMITTEFVLIVTLMHLIRS 356
   | : | | | | | | | | | | | | | | | | | | | | | | | | | |

RESULT 11
Q9SDN3 PRELIMINARY: PRT: 306 AA.
AC 09SDN3:
DB 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, last sequence update)
DR 01-DEC-2001 (TREMBLrel. 19, last annotation update)
DE 1-ACYL-SN-GLYCEROL-3-PHOSPHATE ACYLTRANSFERASE.
OS Prunus dulcis (Almond) (Prunus amygdalus).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eucotsids I; Rosales; Rosaceae; Amygdaloideae; Prunus.
OX NCBI_TaxID=3755;
RN [1]
RP SEQUENCE FROM N.A.
RA Campalans A., Pages M., Messeguier R.;
RT "Identification of differentially expressed genes during dehydration
RT in almond (Prunus amygdalus) using the CDNA-AFlip technique.";
RL Submitted (DEC-1999) to the EMBL/Genbank/DBJ databases.
DR EMBL: AF213937; AAF20003.1;
DR InterPro: IPR002123; Acyltransferase.
DR InterPro: IPR003015; HLH_Myc.
DR Pfam: PF01553; Acyltransferase; 1.
DR ProSITE: PS00038; HELIX_LOOP_HELIX; UNKNOWN_1.
DR Transferrase; Acyltransferase.
KW SEQUENCE 306 AA; 33993 MW; CBD3449625908D CRC64;

Query Match 16.3%; Score 305.5; DB 10; Length 306;
Best Local Similarity 31.9%; Pred. No. 3.3e-20;
Matches 80; Conservative 55; Mismatches 101; Indels 15; Gaps 7;

QY 73 KENIYLANHOSYDVIVADILAIRQNALGHVRYLKEGKLWPLYG--CYFAOHGSIYV 130
   | : | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 3 KHALVISHRSDDIMLVGWVLAQNSGGLSSVAMKRSKFLPYIGSMWSEY--LFL 60
   | : | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 131 KRSKAFNEKEMRNKLSQSYVAGCPMYLVIFESGRYRNPQCTVLSASQAFNAQRGLAVL 190
   | : | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 61 ESMMAKDEGTLKSGVRLKDPQPFMLALFVETGRT--QALLL-AAQGYAATGTPVR 117
   | : | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 191 HVLPRIKATHVAFDCMKNYLAIYDVTVYEGSKDGGORRESPTTEFLCKECPKIHIH 250
   | : | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 118 NVLIRPTKGFVAVSQMSRFAIYDVTAIP-----KSSAPPTMLRLEGRPSVAVH 171
   | : | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 251 IDRIDKVDPEDEHMRMLHERFEIKDKMLEFYSPDPERKRRPPGKSVNSKLSIK 309
   | : | | | | | | | | | | | | | | | | | | | | | | | | | |

DB 172 IRRHVRDLPEDEAVAQWCKDIEFAKDALLDKHTVEQFGDQQLKVTGRPLKS-LLVYT 230
   | : | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 310 TLPSMLISGL 320
   | : | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 231 ANACLLITLAL 241
   | : | | | | | | | | | | | | | | | | | | | | | | | | | |

RESULT 12
Q9VVS1 PRELIMINARY: PRT: 386 AA.
AC 09VVS1: Q9VVS0:
DB 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, last sequence update)
DR 01-JUN-2001 (TREMBLrel. 17, last annotation update)
DE CG4729 PROTEIN.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A. (LONG AND SHORT ISOFORMS).
RC STRAIN-BERKELEY;
RC MEDLINE=20196006; PubMed=10731132.
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galie R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q.X., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazer R.G., Champagne M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abiril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baltwin D.,
RA Baller R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borkova C., Botchan M.R., Bouck J., Brokstein P., Brotlier P.,
RA Burks K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferrara C., Ferreira S., Fleischmann W.,
RA Foster C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwan C.,
RA Jaitani B., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Mishina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Paclab J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Sidenkiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svitskas R., Tector C., Turner R., Ventle E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wasserman D.A., Weinstock G.M., Weissbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster.";
RL Science 287:2185-2195(2000).
-1- ALTERNATIVE PRODUCTS: 2 ISOFORMS; A LONG FORM (SHOWN HERE) AND A
CC SHORT FORM; ARE PRODUCED BY ALTERNATIVE SPLICING.
DB EMBL: AE003527; AAF9471.1;
DB EMBL: AE003527; AAF9472.1;
DB FlyBase: FBgn0036623; CG4729.
DR InterPro: IPR002123; Acyltransferase.
DR Pfam: PF01553; Acyltransferase; 1.
KW Alternative splicing; Hypothetical protein.
FT VARSPLIC 1 243 MISSING (IN SHORT ISOFORM).
SQ SEQUENCE 386 AA; 44397 MW; 1E94370E971C35D5 CRC64;

```

```

Query Match      14.6%; Score 272.5; DB 5; Length 386;
Best Local Similarity 26.6%; Pred. No. 5.6e-17;
Matches 72; Conservative 63; Mismatches 105; Indels 31; Gaps 9;

QY 35 FYQALDEL-----YCYQ--SMVLPFFENYTGQIILYGLDP-----KKNENIYLAN 81
      | : | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Dp 38 FIKPIDKRLFKRLMYACYSLYSQILFVSDWYAGSKMTYMDKEDFEKHAGSEHLLMN 97
      | : | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 82 HQSTVDWIVADILAIRONALGHVRVYLKGLKMLPL--YGCFAOHGCIYKRSKFNEMK 139
      | : | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Dp 98 HKYELIDWLNMGWICKELGVLCNCKAYARKAIRYPILIGWMLAEF--VFLRNRPDOKT 155
      | : | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 140 EMRNKQ--SYVDAGTDMVLIPEEGTRNPEOTKVLASQAFAORGLAVLKHLTPR 196
      | : | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Dp 156 IITBQLKVVFSYPD--FTWILLNAEGTRFTPAKHE--ASVKFAOEGMYVLKHLHLPR 209
      | : | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 197 IKATHVAFDCMKNYLDAYIVTVYVEGDKGQRRSEPTMTFELCKECPKIHIDRIDK 256
      | : | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Dp 210 TKGFVSLAPIRGLCPVYIYDINLAVRPD-----KTPATMLSLHGKSVEPHLMRRRLPL 264
      | : | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 257 KDVEPEOEHMRWLHERFEIKDKMLIEYES 287
      | : | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Dp 265 EOVPEDEKEAAMWLONLFEVKDKIIDSFLT 295
      | : | | | | | | | | | | | | | | | | | | | | | | | | | | | |

RESULT 13
0924S1 PRELIMINARY; PRT; 378 AA.
AC 0924S1;
DT 01-DEC-2001 (Tremblrel. 19, created)
DT 01-DEC-2001 (Tremblrel. 19, last sequence update)
DE 01-DEC-2001 (Tremblrel. 19, last annotation update)
DE LYSOPHOSPHATIDIC ACID ACYLTRANSFERASE-DELTA.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OC NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RA Li W., Suzuki T.;
RT "Rattus norvegicus mRNA for lysophosphatidic acid acyltransferase-
RT delta, complete cds."
RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB067572; BAB62290.1;
KW Transferase; Acyltransferase.
SQ SEQUENCE 378 AA; 43794 MW; 389AA01B7327AE2B CRC64;

```

```

Query Match      14.3%; Score 267; DB 5; Length 380;
Best Local Similarity 25.9%; Pred. No. 1.8e-16;
Matches 83; Conservative 69; Mismatches 123; Indels 46; Gaps 9;

QY 25 RLISAFLLPARYQALDRILCYQSMVLPFFENYTGQIILYGLDPK-----KENIY 78
      | : | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Dp 43 RSLAYLHYFSYCI-----LVCV-----AEWYAGSKRLRYID--PQDEKFFGKEHGIL 89
      | : | | | | | | | | | | | | | | | | | | | | | | | | | | | |

RESULT 14
09V49 PRELIMINARY; PRT; 380 AA.
AC 09V49;
DT 01-MAY-2000 (Tremblrel. 13, created)
DT 01-MAY-2000 (Tremblrel. 13, last sequence update)
DT 01-JUN-2001 (Tremblrel. 17, last annotation update)
DE CG4753 PROTEIN.
GN CG4753.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-BERKELEY;
RC MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazey R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt J., Nelson C.R., Miklos G.L.G.,
RA Abril J.F., Abmayyan A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolintsov S.,
RA Borokova D., Botchan M.R., Bouck J., Brockstein P., Brotler P.,
RA Burris K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Dudin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Foster C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Iodis C.,
RA Jatala M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lal Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Paclet J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirskaas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wasserman D.A., Weinstock G.M., Weissbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zhang X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster."
RL Science 287:2185-2195(2000).
DR EMBL; AE003527; AAF9473.1;
DR FlyBase; FBgn0036622; CG4753.
DR InterPro; IPR002123; Acyltransferase.
DR Pfam; PF01553; Acyltransferase; 1.
SQ SEQUENCE 380 AA; 43482 MW; EA240947566A841 CRC64;

```

[illegible]

```

Db      286 FT-----KDFESDLEVHQINRPILVI-----IMGLVF 318
               | | | | | | | | | |
QY      341 GTLGLVWFIKA 353
               | | | | | | | | | |
Db      319 GGFKLQWLSTVA 331

```

Search completed: August 28, 2002, 11:16:55  
Job time: 559 sec

RESULT 15			
ID	09SYCB	PRELIMINARY;	PRT; 376 AA.
AC	09SYCB;		
DT	01-MAY-2000 (TREMBLrel. 13, Created)		
DT	01-MAY-2000 (TREMBLrel. 13, Last sequence update)		
DT	01-JUN-2001 (TREMBLrel. 17, Last annotation update)		
DE	PUTATIVE ACYL-COA:1-ACYLGLYCEROL-3-PHOSPHATE ACYLTRANSFERASE.		
GN	FLM15.12.		
OS	Arabidopsis thaliana (Mouse-ear cress).		
OC	Eumariophyta; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;		
OC	Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;		
OC	eucotsids II; Brassicales; Brassicaceae; Arabidopsis.		
OX	NCBI_TaxID=3702;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RA	Federpietl N.A., Palm C.J., Conway A.B., Conn L., Hansen N.F.,		
RA	Altai A., Araujo R., Huizar L., Rowley D., Buehler E., Dunn P.,		
RA	Gonzalez A., Kremenetskaia I., Kim C., Lenz C., Li J., Liu S.,		
RA	Lucero S., Schwartz J., Shinn P., Tortilim M., Vysotskaia V.S.,		
RA	Walker M., Yu G., Ecker J., Theologis A., Davis R.W.;		
RL	Submitted (MAY-1999) to the EMBL/GenBank/DBJ databases.		
DR	EMBL; AC006085; AACD30638.1;		
DR	InferPro; IPR002123; Acyltransferase.		
DR	InterPro; IPR003015; HLH_MYC.		
DR	Pfam; PF01553; Acyltransferase; 1.		
DR	PROSITE; PS00038; HELIX_LOOP_HELIX; UNKNOWN_1.		
DR	Transferrase; Acyltransferase.		
QO	SEQUENCE 376 AA; 43435 MW; 0B9AE1A37B09ED4D0 CRC64;		

Query Match	14.18;	Score 263;	DB 10;	Length 376;
Best Local Similarity	23.08;	Pred. No. 4.1e-16;		
Matches 72;	Conservative 76;	Mismatches 117;	Indels 48;	Gaps 8

```

QY 51 VLFEFENTGOVILLYGLPK---NKENIYLNHSOTVDIYADIALAIONALGVRY 106
:: : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 57 LIMEFDMACCKINLYDAETLELIGNEHALVLSNHSIDMLLIGWMAARVCCGLSSLA 116
:: : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 107 VLKGLKMLPLYG--CYFAOHGITYVKRSANEKENRKNLQSYVDAGTMYLVPEEPT 164
:: : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 117 IMKKEAKLPITIGSMFSDY--FELERSNAKDEMLTAKGFKRLDEPDMTFMLALEVEGT 174
:: : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 165 RYNPEQFKVLSASQAPAAQRLANLKLVLPRIKATHVADCMKNYLDATYDVTVYEEG 224
:: : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 175 RFTEQEK---LEAADEYASINSLPSRNVLLPRTKGFVSAVSEIRSEFPAIYDCTLVH-- 229
:: : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 225 DDGQQRRESPTMFEFLCKECPKIIHIDRIDKQDVPPEOECHMRMLERFEIKDKMLIEP 284
:: : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 230 ----NQPTPLLMFSGQSSEINLQMRHRKMSLPETPDGIGIACWQODFLITDAQLEKY 285
:: : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 285 YESDPERRRRPECKSVANSKLT---KTLPSMLILSGLTAGMLMTDAGRKLIVNTWYI 340
:: : : : : : : : : : : : : : : : : : : : : : : : : : : : :

```



**THIS PAGE BLANK (USPTO)**



GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 28, 2002, 11:18:44 ; Search time 75.33 Seconds

(without alignments)  
336.717 Million cell updates/sec

Title: US-09-853-526-5

Perfect score: 364

Sequence: 1 MLSLVLTHTSMRLLPSV.....YVNWITGTLGCLWTRIKA 364

Scoring table:

Gapop 60.0 , Gapext 60.0

Searched: 747574 seqs, 111073796 residues

Word size : 0

Total number of hits satisfying chosen parameters: 747574

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database :

A.Geneseq\_032802:\*  
1: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1980.DAT:\*  
2: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1981.DAT:\*  
3: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1982.DAT:\*  
4: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1983.DAT:\*  
5: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1984.DAT:\*  
6: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1985.DAT:\*  
7: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1986.DAT:\*  
8: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1987.DAT:\*  
9: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1988.DAT:\*  
10: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1989.DAT:\*  
11: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1990.DAT:\*  
12: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1991.DAT:\*  
13: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1992.DAT:\*  
14: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1993.DAT:\*  
15: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1994.DAT:\*  
16: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1995.DAT:\*  
17: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1996.DAT:\*  
18: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1997.DAT:\*  
19: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1998.DAT:\*  
20: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA2000.DAT:\*  
21: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA2001.DAT:\*  
22: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA2001.DAT:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	364	100.0	364	20	AAV36729
2	364	100.0	364	22	AAV50128
3	364	100.0	364	22	AAU00665
4	353	97.0	353	22	AAAB8532
5	353	97.0	353	22	AAAB8552
6	352	96.7	352	20	AAV36728
7	290	79.7	353	22	AAV39591
8	269	73.9	269	22	AAE09653
9	263	72.3	372	22	AAW41377
10	252	69.2	353	22	AAAB93595
11	230	63.2	291	20	AAV36744

12	228	62.6	228	20	AAV36740	Truncated pgl prot
13	199	54.7	261	20	AAV36745	pgl splice variant
14	185	50.8	185	20	AAV36753	pgl splice variant
15	185	50.8	315	20	AAV36751	pgl splice variant
16	184	50.5	300	20	AAV36752	pgl splice variant
17	154	42.3	182	20	AAV36750	pgl splice variant
18	116	31.9	238	20	AAV36743	pgl splice variant
19	85	23.4	96	20	AAV36749	pgl splice variant
20	70	19.2	257	22	ABG01360	Novel human diagno
21	70	19.2	980	22	ABG11977	Novel human diagno
22	62	17.0	66	20	AAV36748	pgl splice variant
23	62	17.0	68	20	AAV36747	pgl splice variant
24	62	17.0	77	20	AAV36742	pgl splice variant
25	62	17.0	90	20	AAV36746	pgl splice variant
26	53	14.6	97	22	ABG12868	Novel human diagno
27	53	14.6	1104	22	ABG12107	Novel human diagno
28	32	8.8	354	20	AAV36741	Mouse pgl protein
29	9	2.5	9	20	AAV36734	pgl box 3 motif.
30	9	2.5	157	22	AAV50130	Lysophosphatidic a
31	9	2.5	237	18	AAV55540	H. pylori ORF 12ge
32	9	2.5	240	19	AAV98477	H. pylori ORF 069p
33	9	2.5	242	18	AAV55441	Acyl glycerol tran
34	8	2.2	9	20	AAV36737	Escherichia coli p
35	8	2.2	239	22	ABB52710	Yeast acyltransfer
36	8	2.2	396	21	AAV99484	Drosophila melanog
37	8	2.2	557	22	ABB71470	Cocoonut Ag3PAT cry
38	8	2.2	1820	22	ABB62370	Cocoonut 1-acylglyc
39	7	1.9	12	16	AAV87705	Cocoonut LPAT clon
40	7	1.9	12	20	AAV43149	1-acylglycerol-3-p
41	7	1.9	65	16	AAV87720	Cocoonut LPAT clon
42	7	1.9	65	20	AAV43164	Cocoonut LPAT clon
43	7	1.9	82	16	AAV87718	Cocoonut LPAT clon
44	7	1.9	82	16	AAV87719	1-acylglycerol-3-p
45	7	1.9	82	20	AAV43160	

#### ALIGNMENTS

RESULT 1	
AAV36729	standard; Protein; 364 AA.
AC	AAV36729;
DT	27-SEP-1999 (first entry)
DE	Human pgl protein sequence.
XX	
KW	PGI gene; diallelic marker; human; PSA: PGI-related diallelic marker;
KW	cancer; prostate cancer; diagnosis; therapy; prostate specific antigen.
OS	Homo sapiens.
XX	
PN	MO9932644-A2.
XX	
PD	01-JUL-1999.
XX	
PF	22-DEC-1998; 98WO-1B02133.
XX	
PR	09-SEP-1998; 98US-0099658.
PR	22-DEC-1997; 97US-0996306.
XX	
PA	(GEST ) GENSET.
PI	Blumenfeld M, Bougueleret L, Chumakov I, Cohen D;
DR	WPI: 1999-405178/34.
DR	N-PSDB; AA00870.
XX	
PT	Use of a prostate cancer associated gene and diallelic markers
XX	derived from it

PS Claim 7; Page 190-191; 365bp; English.

XX The invention relates to a mammalian PGI gene and protein, and a set of  
 CC PGI allelic markers. The PGI polynucleotide and allelic markers are  
 CC used in a hybridisation assay, a sequencing assay, or in an  
 CC allele-specific amplification assay for determining the identity of a  
 CC nucleotide at a PGI-related allelic marker. The methods can be used to  
 CC detect and to assess the risk of developing cancer or prostate cancer.  
 CC Early-stage diagnosis of prostate cancer relies on prostate specific  
 CC antigen (PSA) dosage. However, the effectiveness of this is limited due  
 CC to its inability to discriminate between malignant and non-malignant  
 CC affections of the organ. A need exists for both a reliable diagnostic  
 CC procedure which would enable early-stage diagnosis, and for preventative  
 CC and curative treatments of the disease. The PGI gene can be used for  
 CC detection of prostate cancer, and the risk of developing it in the  
 CC future, and can also be used to determine therapies for the disease.

SO Sequence 364 AA:

Query Match 100.0%; Score 364; DB 20; Length 364;  
 Best Local Similarity 100.0%; Pred. No. 0;  
 Matches 364; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MLSTLVHTTSMKRYLPVSVLLGTAPTYVLAKGVMLLSAPLPARYQALDDRLCYQOS 60  
 DB 1 mlslvtlhtysmyrllpsvvlgltpcyvlawgywrlsflparfygaldrrlcyvys 60

QY 61 MVLFFPNYGVOLLYGSDLPKKNENTITVLANOSTVDWIVADTLATRONALGHVRVVK 120  
 DB 61 mvllffnygvqlllygslpknemillylanqstvdwlvadlaltqnalghvrvylk 120

QY 121 EGLKMLPLCYCFPAHQHGIYVKSARKFNEKEMRNKLOSVDAGTPMYLVFFEGSTRNPE 180  
 DB 121 eglkmlplcycfagbqgylvksarkfnekemrnklsqsyvdagtpmylvffegstrnpe 180

QY 181 QTKVLSAQFPAQRCGLAVLKHVLTPIKATPNAFDCMKVLDATYVTVVYEGSKDDGGO 240  
 DB 181 qtkvlsasqfaeqgrlavlkhvtlprlkathvafdcmkvldatlyvtvvyeqkddgsg 240

QY 241 RRESPTTEFLCKECPRIHIDRIDKKDVPDEQEHRRMLHEFEIKDKMLLEFYSPD 300  
 DB 241 rrespttelckecprihidridkdvdeqehmrmlhefeklkdmllefyspd 300

QY 301 PERRRRPGKSVNSKLSIKRTPSMLTSGLTAGMLMTDAGRKLVTWTVYGTLLGCLMW 360  
 DB 301 perrrrpgksvnskslkrtpsmllsgltagmlmtdagrklyvntwtygtllgclmw 360

QY 361 TTKA 364  
 DB 361 tika 364

RESULT 2  
 AAM50128  
 ID AAM50128 standard; Protein; 364 AA.  
 XX  
 AC AAM50128;  
 XX  
 DT 21-DEC-2001 (first entry)  
 XX  
 DE Human acyltransferase 27417.  
 XX  
 XX Acyltransferase 27417; human; cell proliferation; cell migration;  
 KW cell differentiation; ovary cancer; brain cancer; colon cancer;  
 KW lung cancer; tumour; metastasis; sarcoma; carcinoma;  
 KW adenocarcinoma; antitumour; lipid metabolism; diagnosis; therapy.  
 XX  
 OS Homo sapiens.  
 XX  
 FH Key Location/Qualifiers  
 FT Peptide 1..49  
 FT /Label= Signal\_peptide  
 FT

FT Protein 50..364  
 FT /label= Mature\_protein  
 FT Domain 50..320  
 FT /label= N-terminal\_domain  
 FT Domain 321..337  
 FT /label= Transmembrane\_domain  
 FT Domain 338..364  
 FT /label= C-terminal\_domain  
 FT Domain 71..363  
 FT /label= Acyltransferase\_domain  
 FT Modified-site 68..71  
 FT /note= "Asn is N-glycosylated"  
 FT Modified-site 241..244  
 FT /note= "O-phosphorylated by CAMP- and  
 FT CGMP-dependent protein kinase"  
 FT Modified-site 11..13  
 FT /note= "O-phosphorylated by protein kinase C"  
 FT Modified-site 144..146  
 FT /note= "O-phosphorylated by protein kinase C"  
 FT Modified-site 205..207  
 FT /note= "O-phosphorylated by protein kinase C"  
 FT Modified-site 317..319  
 FT /note= "O-phosphorylated by protein kinase C"  
 FT Modified-site 361..363  
 FT /note= "O-phosphorylated by protein kinase C"  
 FT Modified-site 95..98  
 FT /note= "O-phosphorylated by casein kinase II"  
 FT Modified-site 158..161  
 FT /note= "O-phosphorylated by casein kinase II"  
 FT Modified-site 246..249  
 FT /note= "O-phosphorylated by casein kinase II"  
 FT Modified-site 82..89  
 FT /note= "O-phosphorylated by tyrosine kinase"  
 FT Modified-site 219..226  
 FT /note= "O-phosphorylated by tyrosine kinase"  
 FT Modified-site 23..28  
 FT /note= "N-myristoylated"  
 FT Modified-site 130..135  
 FT /note= "N-myristoylated"  
 FT Modified-site 330..335  
 FT /note= "N-myristoylated"  
 FT Modified-site 352..357  
 FT /note= "N-myristoylated"  
 FT Modified-site 230..343  
 FT /note= "N-myristoylated"  
 FT Modified-site 105..126  
 FT /note= "amidation"  
 FT Peptide  
 FT /note= "predicted leucine zipper"  
 FT  
 PN MO200173051-A2.  
 XX  
 PD 04-OCT-2001.  
 XX  
 PD 26-MAR-2001; 2001WC-US09633.  
 XX  
 PF 24-MAR-2000; 2000US-192092P.  
 XX  
 PR (MILL-) MILLENNIUM PHARM INC.  
 XX  
 PA  
 XX  
 PI  
 XX  
 PI  
 XX  
 DR WPI; 2001-626267/72.  
 DR N-PSDB; AAH27094.  
 XX  
 XX New polynucleotide encodes a polypeptide for the diagnosis and  
 PT treatment of cancer and to alter cellular lipid metabolism comprises a  
 PT polynucleotide encoding human acyltransferases designated 46743 and  
 PT 27417 -  
 XX  
 XX Claim 9; Fig 8A-B; 150bp; English.  
 PS  
 CC The present sequence is that of a novel human protein, termed  
 CC 27417, which shows the structural characteristics of members of the  
 CC acyltransferase family, including an acyltransferase domain. The

CC invention provides novel acyltransferase 27417 nucleic acids and  
 CC polypeptides, as well as methods for detecting their presence, and  
 CC methods for screening for compounds that modulate their expression  
 CC or activity. Such compounds can be used to treat conditions  
 CC related to aberrant activity or expression of 27417 protein or  
 CC nucleic acid, such as conditions involving deficient cellular  
 CC proliferation, migration and/or differentiation. Also provided are  
 CC methods for inhibiting the proliferation or migration, or inducing  
 CC the killing, of a 27417-expressing cell, e.g. a hyperproliferative  
 CC and/or metastatic cell, by contact with a compound that modulates  
 CC the activity or expression of the 27417 protein or nucleic acid.  
 CC The 27417-expressing cell is found in the healthy or diseased  
 CC heart, blood vessels, kidney, skeletal muscle, brain or liver, or  
 CC especially in a solid tumour, a soft tissue tumour or a metastatic  
 CC lesion, a sarcoma, a carcinoma or an adenocarcinoma, and in  
 CC particular in a hyperproliferative and/or metastatic cell found in  
 CC ovarian, brain, colon or lung cancer. The compounds may also be  
 CC used to modulate lipid metabolism in a 27417-expressing cell.  
 CC Methods of disease diagnosis, e.g. by determining the  
 CC presence of a genetic alteration in a 27417 polypeptide, and for  
 CC evaluating the efficacy of a treatment of a disorder, are also  
 CC provided.

XX Sequence 364 AA;

Query Match 100.0%; Score 364; DB 22; Length 364;

Best Local Similarity 100.0%; Pred. No. 0; Mismatches 0; Indels 0; Gaps 0;

Matches 364; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MLISLVHTYSMRYLLPSVVLGTAPTYVLAMGVMRLSAFLPARYQALDRLXCVYGS 60  
 DB 1 mlislvlhysmryllpsvvlgtaptlylwgvmrllsaflparfygaldtrlycvygs 60  
 QY 61 MYLFFENYTGVOILLXGDIPLKKNENIYLANHSTVDWIVADIARONALGHVRYLK 120  
 DB 61 mylffenytyvgillxgdlpkkneniylanhstvdwivadiarlqnalghvrylk 120  
 QY 121 EGLKMLPIVGCYPAOHGCIYVKRSAPFNEKEMNKLOSVDAGTPMYLVTFPGSTRYNP 180  
 DB 121 eglkmlpivgcypaohgciyvkrsakfnekemnklosyvdagctpmylvtfpgstrynpe 180  
 QY 181 QTKVLSAQAFAAQRGLAVLKHVLTPIRIKATHVAFDCKMNYLDAIDVTVVEGKDDGQ 240  
 DB 181 qtkvlsasqafaagrglavlkhlvtprikathvafdcmknyldaivdvtevyegkddggq 240  
 QY 241 RRSPTMTFELCKECPKIHIDRIDKDVPEQEHMRMLHREFEIKDKMLIEFYESP 300  
 DB 241 rrespmtfclckecpkihidridkdvpeeqehmrmlherfeikdkmliefespd 300  
 QY 301 PERRRRPCKSVNSKISIKKTLPSMLILSGLTAGMLMTDAGRKLYNWTWYTGTLGCIWY 360  
 DB 301 perrrkfpkgsvnsksikklpsmllsgltagmimtdagrkllynwtwlytgclgw 360  
 QY 361 TTKA 364  
 DB 361 tika 364

RESULT 3  
 AAU00665  
 ID AAU00665 standard; Protein; 364 AA.

XX AAU00665;

XX 07-SEP-2001 (first entry)

DE Human lysophosphatidic acid acyltransferase isoform LPAAT-epsilon.

KX Lysophosphatidic acid acyltransferase epsilon; LPAAT-epsilon; LPA; PA;  
 KW Lysophosphatidic acid; phosphatidic acid; acylation; cellular activation;  
 KW phospholipid signaling; mitogenesis; inflammation; autoimmune disease;  
 KW oncology; cancer; obesity; gene therapy.

XX Homo sapiens.  
 OS  
 XX  
 PN WO200134782-A1.  
 XX  
 PD 17-MAY-2001.  
 XX  
 PF 02-NOV-2000; 2000WO-US30193.  
 XX  
 PR 09-NOV-1999; 99US-0436919.  
 XX  
 PA (CELL-) CELL THERAPEUTICS INC.  
 PI Leung DW;  
 XX  
 DR WPI; 2001-335920/35.  
 DR N-PSDB; AAS00649.  
 PT Novel isolated human isoform of lysophosphatidic acid  
 PT acyltransferase-epsilon useful for diagnostic, therapeutic and  
 PT screening purposes -  
 XX  
 PS Claim 1; Fig 1; 48pp; English.

The sequence represents a human lysophosphatidic acid acyltransferase  
 CC (LPAAT) isoform, LPAAT-epsilon. LPAAT catalyses the acylation of  
 CC lysophosphatidic acid (LPA) to phosphatidic acid (PA). LPA and PA have  
 CC been identified as phospholipid signalling molecules that affect a wide  
 CC range of biological responses. PA is involved in cellular activation and  
 CC mitogenesis. Compounds that block PA generation and hence diminish lipid  
 CC biosynthesis and the signal involved in cell activation are of  
 CC therapeutic interest in the areas of inflammation and oncology (e.g.  
 CC autoimmune diseases and cancer) as well as obesity treatment.  
 CC LPAAT-epsilon and its corresponding DNA can be used in screening assays  
 CC to detect agents that stimulate or inhibit the activity of LPAAT and,  
 CC therefore, PA. The DNA is useful in tests to detect the presence or  
 CC expression of LPAAT-epsilon in relation to certain diseases and  
 CC conditions, and in disease prevention and treatment. The sequences of the  
 CC invention are also useful for diagnosis of diseases and conditions in  
 CC which the expression of LPAAT enzyme is abnormal.

XX Sequence 364 AA;

Query Match 100.0%; Score 364; DB 22; Length 364;

Best Local Similarity 100.0%; Pred. No. 0; Mismatches 0; Indels 0; Gaps 0;

Matches 364; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MLISLVHTYSMRYLLPSVVLGTAPTYVLAMGVMRLSAFLPARYQALDRLXCVYGS 60  
 DB 1 mlislvlhysmryllpsvvlgtaptlylwgvmrllsaflparfygaldtrlycvygs 60  
 QY 61 MYLFFENYTGVOILLXGDIPLKKNENIYLANHSTVDWIVADIARONALGHVRYLK 120  
 DB 61 mylffenytyvgillxgdlpkkneniylanhstvdwivadiarlqnalghvrylk 120  
 QY 121 EGLKMLPIVGCYPAOHGCIYVKRSAPFNEKEMNKLOSVDAGTPMYLVTFPGSTRYNP 180  
 DB 121 eglkmlpivgcypaohgciyvkrsakfnekemnklosyvdagctpmylvtfpgstrynpe 180  
 QY 181 QTKVLSAQAFAAQRGLAVLKHVLTPIRIKATHVAFDCKMNYLDAIDVTVVEGKDDGQ 240  
 DB 181 qtkvlsasqafaagrglavlkhlvtprikathvafdcmknyldaivdvtevyegkddggq 240  
 QY 241 RRSPTMTFELCKECPKIHIDRIDKDVPEQEHMRMLHREFEIKDKMLIEFYESP 300  
 DB 241 rrespmtfclckecpkihidridkdvpeeqehmrmlherfeikdkmliefespd 300  
 QY 301 PERRRRPCKSVNSKISIKKTLPSMLILSGLTAGMLMTDAGRKLYNWTWYTGTLGCIWY 360  
 DB 301 perrrkfpkgsvnsksikklpsmllsgltagmimtdagrkllynwtwlytgclgw 360  
 QY 361 TTKA 364

```
|||||
Db      361 tika 364

RESULT  4
AAB85532 standard; protein; 353 AA.
XX
AC      AAB85532:
XX
DT      25-SEP-2001 (first entry)
XX
DE      Human secreted protein (clone id HLTH084).
XX
KW      Secreted protein; immunosuppressive; antiarthritic; antirheumatic;
KW      antiproliferative; cytostatic; cardiact; vasotropic; cerebroprotective;
KW      neurotropic; neuroprotective; antibacterial; virucide; fungicide; human;
KW      opthalmological; gene therapy.
XX
OS      Homo sapiens.
XX
PN      WO200155430-A1.
XX
PD      02-AUG-2001.
XX
PF      17-JAN-2001: 2001WO-US01431.
XX
PR      31-JAN-2000: 2000US-0179065.
PR      04-FEB-2000: 2000US-0180628.
PR      12-SEP-2000: 2000US-0231968.
XX
PA      (HUMA-) HUMAN GENOME SCI INC.
XX
PI      Rosen CA, Komatsoulis GA, Baker KP, Birse CE, Soppet DR, Olsen HS;
PI      Moore PA, Wei P, Ebner R, Duan DR, Shi Y, Choi GH, Fiscella M;
PI      Ni J, Ruben SM, Barash SC;
XX
DR      WPI: 2001-476220/51.
DR      N-PSTDB: AAAH46942.
XX
PT      17 Isolated nucleic acid molecules encoding human secreted proteins,
PT      used to preventing, treating or ameliorating a medical condition -
XX
PS      Claim 11: Page 450-451; 482pp; English.
XX
CC      The invention provides novel human secreted proteins and polynucleotides
CC      encoding them. The secreted proteins can be expressed by standard
CC      recombinant methodology. The secreted proteins and polynucleotides are
CC      used to prevent, treat or ameliorate a medical condition in e.g. humans,
CC      mice, rabbits, goats, horses, cats, dogs, chickens or sheep. They can
CC      also be used in diagnosing a pathological condition. The antibodies to
CC      the proteins can also be used in alleviating symptoms associated with the
CC      disorders and in diagnostic immunoassays e.g. radioimmunoassays or enzyme
CC      linked immunosorbent assays (ELISA). Disorders which are diagnosed or
CC      treated include autoimmune diseases e.g. rheumatoid arthritis,
CC      hyperproliferative disorders e.g. neoplasms of the breast or liver,
CC      cardiovascular disorders e.g. cardiac arrest, cerebrovascular disorders
CC      e.g. cerebral ischemia, anglogenesis, nervous system disorders e.g.
CC      Alzheimer's disease, infections caused by bacteria, viruses and fungi and
CC      ocular disorders e.g. corneal infection. The polypeptides can also be
CC      used to aid wound healing and epithelial cell proliferation, to prevent
CC      skin aging due to sunburn, to maintain organs before transplantation, for
CC      supporting cell culture of primary tissues, to regenerate tissues and in
CC      chemotaxis. The polypeptides can also be used as a food additive or
CC      preservative to increase or decrease storage capabilities. The present
CC      sequence represents a human secreted protein.
XX
SQ      Sequence      353 AA;
```

```
QY      12 MRYLPSVVLGTAPTYVLVANGVWRLLSAFLPARFYOALDDRLKCYQSWLFFPENYTG 71
Db      1 mryllpsvvlgltpatyvllangvwrllsalfparfygaldrlkcyqswvlfffenyig 60
QY      72 VOILLGDDLPKNENITYLANNHSTVDWIVADILATRONALGHVRVLRKGLKWLPGC 131
Db      61 vqlllgddlpknennilylannhgstvdwivadilaltrqnalghvrvlkeglkwlpygc 120
QY      132 YFAOHGSIYVRSKAKENKEEMRNKLOSVDAGTPMYLVFPEEGTRYNPEQTKVLSAQAF 191
Db      121 yfaohgsllyvkrskakfnekemrnklqsyvdagtpmylvlfpeegtrynpeqtkvlssasqaf 180
QY      192 AAQRGVLAVLKHVLTTPRKATHVAFDCMKNLDAIYDVTVYVECKDGGQRRESPTMTTEL 251
Db      181 aaqrgvlavlkhvltprikathvafdcmknyldaiydvttvyvegkddgqrresptmtel 240
QY      252 CKECPRKHNIHIDRKDDVEEEOHMRWRHFEFEIKDKMLIFEYSPDERKRRPGKS 311
Db      241 ckeoprkhnihidrkddveeoeohmrwrhfefelkdcmlliefespderrkrrpgks 300
QY      312 VNSKLSIKKTLPSMLILSGLTAGMLTDAGRKLYVTWTVYGTLLGLLWYTIKA 364
Db      301 vnskslkkcltpsmllsgltagmltdagrkllyvntwlygtllgclwvrika 353

RESULT  5
AAB85552 standard; protein; 353 AA.
XX
AC      AAB85552:
XX
DT      25-SEP-2001 (first entry)
XX
DE      Human secreted protein (clone id HSLIA81).
XX
KW      Secreted protein; immunosuppressive; antiarthritic; antirheumatic;
KW      antiproliferative; cytostatic; cardiact; vasotropic; cerebroprotective;
KW      neurotropic; neuroprotective; antibacterial; virucide; fungicide; human;
KW      opthalmological; gene therapy.
XX
OS      Homo sapiens.
XX
PN      WO200155430-A1.
XX
PD      02-AUG-2001.
XX
PF      17-JAN-2001: 2001WO-US01431.
XX
PR      31-JAN-2000: 2000US-0179065.
PR      04-FEB-2000: 2000US-0180628.
PR      12-SEP-2000: 2000US-0231968.
XX
PA      (HUMA-) HUMAN GENOME SCI INC.
XX
PI      Rosen CA, Komatsoulis GA, Baker KP, Birse CE, Soppet DR, Olsen HS;
PI      Moore PA, Wei P, Ebner R, Duan DR, Shi Y, Choi GH, Fiscella M;
PI      Ni J, Ruben SM, Barash SC;
XX
DR      WPI: 2001-476220/51.
DR      N-PSTDB: AAAH46962.
XX
PT      17 Isolated nucleic acid molecules encoding human secreted proteins,
PT      used to preventing, treating or ameliorating a medical condition -
XX
PS      Claim 11: Page 465-466; 482pp; English.
XX
CC      The invention provides novel human secreted proteins and polynucleotides
CC      encoding them. The secreted proteins can be expressed by standard
CC      recombinant methodology. The secreted proteins and polynucleotides are
CC      used to prevent, treat or ameliorate a medical condition in e.g. humans,
CC      mice, rabbits, goats, horses, cats, dogs, chickens or sheep. They can
CC      also be used in diagnosing a pathological condition. The antibodies to
```

CC the proteins can also be used in alleviating symptoms associated with the  
CC disorders and in diagnostic immunoassays e.g. radioimmunoassays or enzyme  
CC linked immunosorbent assays (ELISA). Disorders which are diagnosed or  
CC treated include autoimmune diseases e.g. rheumatoid arthritis,  
CC hyperproliferative disorders e.g. neoplasms of the breast or liver,  
CC cardiovascular disorders e.g. cardiac arrest, cerebrovascular disorders  
CC e.g. cerebral ischemia, angiodenesis, nervous system disorders e.g.  
CC Alzheimer's disease, infections caused by bacteria, viruses and fungi and  
CC ocular disorders e.g. corneal infection. The polypeptides can also be  
CC used to aid wound healing and epithelial cell proliferation, to prevent  
CC skin aging due to sunburn, to maintain organs before transplantation, for  
CC supporting cell culture of primary tissues, to regenerate tissues and in  
CC chemotaxis. The polypeptides can also be used as a food additive or  
CC preservative to increase or decrease storage capabilities. The present  
CC sequence represents a human secreted protein.

XX Sequence 353 AA;

Query Match 97.0%; Score 353; DB 22; Length 353;  
Best local Similarity 100.0%; Pred. No. 0;  
Matches 353; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 12 MRLTSLVLLGTPPTVYVAMGWRLISAFLEPRFYQALDDRLCYQSWLFEFFENYTG 71  
DB 1 MYLLPSVLLGTAPYVLAHWYVWLLSALFIPARFYGALDDRLCYQSWLFEFFENYTG 60

QY 72 VOLLVGDLPKKNKENTIIYLANHSTVDWIVADLIRONALGHVRYLKGKWLPLYGC 131  
DB 61 VQILLYGDLPKKNKENTIIYLANHSTVDWIVADLIRONALGHVRYLKGKWLPLYGC 120

QY 132 YFAOHGIIYKRSRAKFEKEMRNKLSYVDAGTPMYLVFPESTRYNEPTKYLASQAF 191  
DB 121 YFQHGIIYKRSRAKFEKEMRNKLSYVDAGTPMYLVFPESTRYNEPTKYLASQAF 180

QY 192 AAGRGAVLKHNVLTPRIKATHVAFDCMKNYLAIYDVTVYVEGKDDGGGRRSPMTTEFL 251  
DB 181 AAGRGAVLKHNVLTPRIKATHVAFDCMKNYLAIYDVTVYVEGKDDGGGRRSPMTTEFL 240

QY 252 CKRCPRKHHIHDIDKDDVEEBOEHMRRLHEFEIKDKMLIEFYSPDERKKRPFGKS 311  
DB 241 CKRCPRKHHIHDIDKDDVEEBOEHMRRLHEFEIKDKMLIEFYSPDERKKRPFGKS 300

QY 312 VNSKLSIKKTLPMSLLISGLTAGMLMTDAGRKLIVNTWIVYGTLLGCLWVTRIK 364  
DB 301 VNSKLSIKKTLPMSLLISGLTAGMLMTDAGRKLIVNTWIVYGTLLGCLWVTRIK 353

RESULT 6

AA36728 ID AAY36728 standard; Protein; 352 AA.

AC AAY36728;

DT 27-SEP-1999 (first entry)

DE Human PGI protein sequence.

KW PGI gene; biallelic marker; human; PSA; PGI-related biallelic marker;  
KW cancer; prostate cancer; diagnosis; therapy; prostate specific antigen.

OS Homo sapiens.

PN WO932644-A2.

PD 01-JUL-1999.

PF 22-DEC-1998; 98WO-1B02133.

PR 09-SEP-1998; 98US-0099658.

PR 22-DEC-1997; 97US-0996506.

PA (GEST ) GENSET.

XX Blumenfeld M, Bougueleret L, Chumakov I, Cohen D;  
PI MPI: 1999-405178/34.  
DR N-PSDB; AAZ00872.

XX Use of a prostate cancer associated gene and biallelic markers  
XX derived from it  
PS Claim 7; Page 190; 385pp; English.

CC The invention relates to a mammalian PGI gene and protein, and a set of  
CC PGI biallelic markers. The PGI polynucleotide and biallelic markers are  
CC used in a hybridization assay, a sequencing assay, or in an  
CC allele-specific amplification assay for determining the identity of a  
CC nucleotide at a PGI-related biallelic marker. The methods can be used to  
CC detect and to assess the risk of developing cancer or prostate cancer.  
CC Early-stage diagnosis of prostate cancer relies on prostate specific  
CC antigen (PSA) dosage. However, the effectiveness of this is limited due  
CC to its inability to discriminate between malignant and non-malignant  
CC affections of the organ. A need exists for both a reliable diagnostic  
CC procedure which would enable early-stage diagnosis, and for preventative  
CC and curative treatments of the disease. The PGI gene can be used for  
CC detection of prostate cancer, and the risk of developing it in the  
CC future, and can also be used to determine therapies for the disease.

XX Sequence 352 AA;

Query Match 96.7%; Score 352; DB 20; Length 352;  
Best local Similarity 100.0%; Pred. No. 0;  
Matches 352; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 12 MRLTSLVLLGTPPTVYVAMGWRLISAFLEPRFYQALDDRLCYQSWLFEFFENYTG 71  
DB 1 MYLLPSVLLGTAPYVLAHWYVWLLSALFIPARFYGALDDRLCYQSWLFEFFENYTG 60

QY 72 VOLLVGDLPKKNKENTIIYLANHSTVDWIVADLIRONALGHVRYLKGKWLPLYGC 131  
DB 61 VQILLYGDLPKKNKENTIIYLANHSTVDWIVADLIRONALGHVRYLKGKWLPLYGC 120

QY 132 YFAOHGIIYKRSRAKFEKEMRNKLSYVDAGTPMYLVFPESTRYNEPTKYLASQAF 191  
DB 121 YFQHGIIYKRSRAKFEKEMRNKLSYVDAGTPMYLVFPESTRYNEPTKYLASQAF 180

QY 192 AAGRGAVLKHNVLTPRIKATHVAFDCMKNYLAIYDVTVYVEGKDDGGGRRSPMTTEFL 251  
DB 181 AAGRGAVLKHNVLTPRIKATHVAFDCMKNYLAIYDVTVYVEGKDDGGGRRSPMTTEFL 240

QY 252 CKRCPRKHHIHDIDKDDVEEBOEHMRRLHEFEIKDKMLIEFYSPDERKKRPFGKS 311  
DB 241 CKRCPRKHHIHDIDKDDVEEBOEHMRRLHEFEIKDKMLIEFYSPDERKKRPFGKS 300

QY 312 VNSKLSIKKTLPMSLLISGLTAGMLMTDAGRKLIVNTWIVYGTLLGCLWVTRIK 363  
DB 301 VNSKLSIKKTLPMSLLISGLTAGMLMTDAGRKLIVNTWIVYGTLLGCLWVTRIK 352

RESULT 7

AA39591 ID AAM39591 standard; Protein; 353 AA.

AC AAM39591;

DT 22-OCT-2001 (first entry)

DE Human polypeptide SEQ ID NO 2736.

KW Human; nootropic; immunosuppressant; cytostatic; gene therapy; cancer;  
KW peripheral nervous system; neuropathy; central nervous system; CNS;  
KW Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;  
KW amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;  
KW chemokinetic; thrombolytic; drug screening; arthritis; inflammation;

KW Leukaemia.  
XX Homo sapiens.  
OS  
XX WO200153312-A1.  
XX  
XX 26-JUL-2001.  
XX  
XX 26-DEC-2000; 2000WO-US34263.  
XX  
XX 21-JAN-2000; 2000US-0488725.  
PR 25-APR-2000; 2000US-0552317.  
PR 09-JUL-2000; 2000US-0598042.  
PR 19-JUL-2000; 2000US-0620312.  
PR 03-AUG-2000; 2000US-0653450.  
PR 14-SEP-2000; 2000US-0662191.  
PR 19-OCT-2000; 2000US-0693036.  
PR 29-NOV-2000; 2000US-0727344.  
XX  
XX (HYSE-) HYSEQ INC.  
XX  
XX Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;  
PI Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J;  
PI Zhao QA, Zhou P, Goodrich R, Drmanac RT;  
XX  
XX WPI: 2001-442253/47.  
DR N-PSDB; AAI58747.  
XX  
XX Novel nucleic acids and polypeptides, useful for treating disorders  
PT such as central nervous system injuries -  
XX  
XX Example 4; SEQ ID NO 2736; 10078bp; English.  
XX  
XX The invention relates to human nucleic acids (AA157798-AA161369) and  
CC the encoded polypeptides (AA38642-AA42213) with neotropic,  
CC immunosuppressant and cyostatic activity. The polynucleotides are useful  
CC in gene therapy. A composition containing a polypeptide or polynucleotide  
CC of the invention may be used to treat diseases of the peripheral nervous  
CC system, such as peripheral nervous injuries, peripheral neuropathy and  
CC localised neuropathies and central nervous system diseases, such as  
CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic  
CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the  
CC utilisation of the activities such as: Immune system suppression,  
CC Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic  
CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,  
CC assays for receptor activity, arthritis and inflammation, leukaemias and  
CC C.N.S disorders.  
CC Note: The sequence data for this patent did not form part of the printed  
CC specification.  
XX  
XX Sequence 353 AA;  
SQ

Query Match 79.7%; Score 290; DB 22; Length 353;  
Best Local Similarity 100.0%; Pred. No. 9.9e-286;  
Matches 290; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 12 MRYLLPSVLLGTAPTYVLAWGVRLSAFLPARFYQALDRLCYQSMWLEFFENYTG 71  
DB 1 mryllpsvllgtaptyvlawgvrlsaflparfyqaldrlycvysmwlffenytyg 60  
QY 72 VOILLGDLPRKNENITYLANKHSTVDMIVADILAIRONALGHVRYLKESLKLPLPGC 131  
DB 61 vqlllygdldpknkenillylanhgstvdmivadilaqrnalghvrylkeglikwlpkygc 120  
QY 132 YFAOHGIIYVRSKAKNEKEMKRNKLSYVDAGTDMYLVIFPEGTRVPEOTKVISASQAF 191  
DB 121 yfaohggiivvrsaknekekernklsyvdagtpmylvifpegtrvpeotkviasaqaf 180  
QY 192 AAOGLAVLVHVTLPRIKATHVAFDCKKNYLDAIVDTVYVEGDDCGGORESTPTEFL 251  
DB 181 aagtglaavlkhvltprirkat hvaf dckkn yldai vdtvv yegddgggrr estp tlef l 240

QY 252 CKCEPKIHIIIDRIDKQVPEDEHMRRLHERPEIKDKMLIEFYSPDP 301  
DB 241 ckecpk ihidridk qvpe de hmr r l herpe i kdk m l ief ysp dp 290

RESULT 8  
ID AAE09653 standard; Protein; 269 AA.  
XX  
XX AAE09653;  
AC  
XX  
XX 19-NOV-2001 (first entry)  
DE  
XX Human gene 15 encoded lipid metabolism protein HHASB32. SEQ ID NO:47.  
KW Human: lipid metabolism protein; neotropic; neuroprotective; cardiant;  
KW cerebroprotective; hepatotoxic; antidiabetic; ophthalmic; nephrotropic;  
KW immune disorder; autoimmune disease; rheumatoid arthritis; glossitis;  
KW systemic lupus erythematosus; conjunctivitis; inflammatory disorder;  
KW respiratory disorder; asthma; allergy; CNS disorder; Alzheimer's disease;  
KW Parkinson's disease; atherosclerosis; cardiovascular disorder; cancer;  
KW coronary disease; familial hypercholesterolaemia; hyperlipidaemia;  
KW hematoietic disorder; hypolipidaemia; lipidosi; Gaucher's disease;  
KW Tay-sach's disease; mental retardation; gene therapy; antisense therapy.  
XX  
XX Homo sapiens.  
OS  
XX WO200155203-A1.  
XX  
XX 02-AUG-2001.  
PD  
XX  
XX 17-JAN-2001; 2001WO-US01327.  
XX  
XX 31-JAN-2000; 2000US-0179065.  
PR 04-FEB-2000; 2000US-0180628.  
PR 24-FEB-2000; 2000US-0184664.  
PR 02-MAR-2000; 2000US-0186350.  
PR 16-MAR-2000; 2000US-0189674.  
PR 17-MAR-2000; 2000US-0190076.  
PR 18-APR-2000; 2000US-0198123.  
PR 19-MAY-2000; 2000US-0205515.  
PR 07-JUN-2000; 2000US-0209467.  
PR 28-JUN-2000; 2000US-0214886.  
PR 30-JUN-2000; 2000US-0215135.  
PR 07-JUL-2000; 2000US-0216647.  
PR 07-JUL-2000; 2000US-0216680.  
PR 11-JUL-2000; 2000US-0217487.  
PR 11-JUL-2000; 2000US-0217496.  
PR 14-JUL-2000; 2000US-0218390.  
PR 26-JUL-2000; 2000US-0220363.  
PR 26-JUL-2000; 2000US-0220964.  
PR 14-AUG-2000; 2000US-0224518.  
PR 14-AUG-2000; 2000US-0224519.  
PR 14-AUG-2000; 2000US-0225213.  
PR 14-AUG-2000; 2000US-0225214.  
PR 14-AUG-2000; 2000US-0225267.  
PR 14-AUG-2000; 2000US-0225268.  
PR 14-AUG-2000; 2000US-0225270.  
PR 14-AUG-2000; 2000US-0225270.  
PR 14-AUG-2000; 2000US-0225447.  
PR 14-AUG-2000; 2000US-0225757.  
PR 14-AUG-2000; 2000US-0225758.  
PR 14-AUG-2000; 2000US-0225759.  
PR 18-AUG-2000; 2000US-0226279.  
PR 22-AUG-2000; 2000US-0226681.  
PR 22-AUG-2000; 2000US-0226688.  
PR 22-AUG-2000; 2000US-0227182.  
PR 23-AUG-2000; 2000US-0227009.  
PR 30-AUG-2000; 2000US-0228924.  
PR 01-SEP-2000; 2000US-0229287.  
PR 01-SEP-2000; 2000US-0229343.  
PR 01-SEP-2000; 2000US-0229344.  
PR 01-SEP-2000; 2000US-0229345.

Query Match	Best Local Similarity	73.9%;	Score 269;	DB 22;	Length 269;
Matches 269;	Conservative	0;	Mismatches	0;	Indels 0;
Gaps	0;				
96	TVDVADIIILIRONALGHVRYLKEGKWTPLKGCYFAOHGIGTYVRSKAFNEKRNK	155			
1	TVQWIVADIIILIRQALGHVRYLKEGKWTPLKGCYFAOHGIGTYVRSKAFNEKRNK	60			
156	LSQVYDAGTTPMYLYVIFPEGRTYRNPEDQTKVLSASQAFNAORGLAVLKHTLPRIKATHVAF	215			
61	IGSYVDAGTTPMYLYVIFPEGRTYRNPEDQTKVLSASQAFNAORGLAVLKHTLPRIKATHVAF	120			

QY 216 DCKNNTDAIYDTVYVEGDDGQRESEPTMEFECKECPKIHIDRIDKDVPEEQE 275  
 DB 121 dcmknnyidaivdyvvyegkddgqrrrespmtelfckeopklnhldridkdvpege 180  
 QY 276 HMRRLHEREFEIKDKMLIEFYESPDPERRRKRPFGKSVNSKLSIKKTLPSMLILISGLTAGM 335  
 DB 181 hmrrrlherfelkdkmliefyespdperkrfrfpgksvnsklsiktlpsmlilsgltagm 240  
 QY 336 LMTDAGRRLVNTWYVYGTLLGCLMWTIKA 364  
 DB 241 lmtdagrklyvntwiygtllgclwltika 269

RESULT 9  
 AAM41377  
 ID AAM41377 standard; Protein: 372 AA.  
 XX  
 AC AAM41377;  
 XX  
 DT 22-OCT-2001 (first entry)  
 XX  
 DE Human polypeptide SEQ ID NO 6308.  
 XX  
 KW Human: neurotrophic; immunosuppressant; cytostatic; gene therapy; cancer;  
 KW peripheral nervous system; neuropathy; central nervous system; CNS;  
 KW Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;  
 KW amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;  
 KW chemokine; thrombolytic; drug screening; arthritis; inflammation;  
 KW leukemia.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200153312-A1.  
 PD 26-JUL-2001.  
 XX  
 PF 26-DEC-2000; 2000WO-US34263.  
 XX  
 PR 21-JAN-2000; 2000US-0488725.  
 PR 25-APR-2000; 2000US-0552317.  
 PR 09-JUL-2000; 2000US-0598042.  
 PR 19-JUL-2000; 2000US-0620312.  
 PR 03-AUG-2000; 2000US-0653450.  
 PR 14-SEP-2000; 2000US-0662191.  
 PR 19-OCT-2000; 2000US-0693036.  
 PR 29-NOV-2000; 2000US-0727344.  
 XX  
 PA (HYSE-) HYSEQ INC.  
 XX  
 PI Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;  
 PI Wang J, Wang Z, Wehman T, Xu C, Xue AJ, Yang Y, Zhang J;  
 PI Zhao QA, Zhou P, Goodrich R, Drmanac RT;  
 XX  
 DR WPI: 2001-442253/47.  
 DR N-PSDB: AA160533.  
 XX  
 PT Novel nucleic acids and polypeptides, useful for treating disorders  
 PT such as central nervous system injuries -  
 XX  
 PS Example 2; SEQ ID NO 6308; 10078bp; English.  
 XX  
 CC The invention relates to human nucleic acids (AA157798-AA161369) and  
 CC the encoded polypeptides (AAM38642-AAM42213) with neurotrophic,  
 CC immunosuppressant and cytostatic activity. The polynucleotides are useful  
 CC in gene therapy. A composition containing a polypeptide or polynucleotide  
 CC of the invention may be used to treat diseases of the peripheral nervous  
 CC system, such as peripheral nervous injuries, peripheral neuropathy and  
 CC localised neuropathies and central nervous system diseases, such as  
 CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic  
 CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the  
 CC utilisation of the activities such as: immune system suppression,  
 CC Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic  
 CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,

CC assays for receptor activity, arthritis and inflammation, leukaemias and  
 CC C.N.S disorders.  
 CC Note: The sequence data for this patent did not form part of the printed  
 CC specification.  
 CC  
 SQ Sequence 372 AA:

Query Match 72.3%; Score 263; DB 22; Length 372;  
 Best Local Similarity 99.7%; Pred. No. 2,7e-258;  
 Matches 363; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MLSLVHTYSMRYLPSVVLGTAFTYVLAWGVRLSAFLPAREYQALDRLYCVYGS 60  
 DB 9 mlslvhtysmryllpsvvlgtaptlyawgvrlsaflparfygalddrlycvygs 68  
 QY 61 MVLFFPENTGYOILLYGDI-PKKNKENTIYLANHOSYVDWIVADILAIROMALGHVRYLK 120  
 DB 69 mvlffpentygqililygdipkknkenliylanhsctvdwlvadilalrqnalghvrylk 128  
 QY 121 EGLKWLPLXGCFYAOHNGIYVKSRAKENEKEMRNKLQSYVDAGTPLYVIFPGSTRYNPE 180  
 DB 129 eglkwlplygwyfaqbgilyvksakfnekemrnklqsyvdagtpmlyvifpgstrynpe 188  
 QY 181 QTVYLSASQAPAAQRCGLAVLKHYLTPRIKATHYAFDCMKNYLDIYDYVYVYGGKDDGQ 240  
 DB 189 qtvylsasqaafaqrglavlkhytlprikatbvaafdomknyldiaydvtyvygkddgq 248  
 QY 241 RRESPWTEFLCKECPKIHIDRIDKDVPEEOEHRRLHBRFEIKDKMLIEFESPD 300  
 DB 249 rrespmtelfckeopklnhldridkdvpeeehmrrlherfelkdkmliefespd 308  
 QY 301 PERRRRPFGKSVNSKLSIKKTLPSMLILSGTAGMLMTDAGRRLVYNTWYVYGTLLGCLWY 360  
 DB 309 perrrrpfpgksvnsklsikktlpsmlilsgltagmlmtdagrklyvntwiygtllgclw 368  
 QY 361 TIRKA 364  
 DB 369 tlrka 372

RESULT 10  
 AAB93595  
 ID AAB93595 standard; Protein: 353 AA.  
 XX  
 AC AAB93595;  
 XX  
 DT 26-JUN-2001 (first entry)  
 XX  
 DE Human protein sequence SEQ ID NO:13028.  
 XX  
 KW Human; primer; detection; diagnosis; antisense therapy; gene therapy.  
 XX  
 OS Homo sapiens.  
 XX  
 PN EP1074617-A2.  
 PD 07-FEB-2001.  
 XX  
 PF 28-JUL-2000; 2000EP-0116126.  
 XX  
 PR 29-JUL-1999; 99JP-0248036.  
 PR 27-AUG-1999; 99JP-0300253.  
 PR 11-JAN-2000; 2000JP-0118776.  
 PR 02-MAY-2000; 2000JP-0183767.  
 PR 09-JUN-2000; 2000JP-0241899.  
 XX  
 PA (HELI-) HELIX RES INST.  
 XX  
 PI Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;  
 PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;  
 XX  
 DR WPI: 2001-318749/34.



XX Primer sets for synthesizing polynucleotides, particularly the 5602  
PT full-length cDNAs defined in the specification, and for the detection  
PT and/or diagnosis of the abnormality of the proteins encoded by the  
PT full-length cDNAs -  
PS Claim 8; SEQ ID 13028; 2537bp + CD ROM; English.  
XX  
XX The present invention describes primer sets for synthesizing 5602  
CC full-length cDNAs defined in the specification. Where a primer set  
CC comprises: (a) an oligo-dT primer and an oligonucleotide complementary  
CC to the complementary strand of a polynucleotide which comprises one of  
CC the 5602 nucleotide sequences defined in the specification, where the  
CC oligonucleotide comprises at least 15 nucleotides; or (b) a combination  
CC of an oligonucleotide comprising a sequence complementary to the  
CC complementary strand of a polynucleotide which comprises a 5'-end  
CC sequence and an oligonucleotide comprising a sequence complementary to a  
CC polynucleotide which comprises a 3'-end sequence, where the  
CC oligonucleotide comprises at least 15 nucleotides and the combination of  
CC the 5'-end sequence/3'-end sequence is selected from those defined in  
CC the specification. The primer sets can be used in antisense therapy and  
CC in gene therapy. The primers are useful for synthesizing polynucleotides,  
CC particularly full-length cDNAs. The primers are also useful for the  
CC detection and/or diagnosis of the abnormality of the proteins encoded by  
CC the full-length cDNAs. The primers allow obtaining of the full-length  
CC cDNAs easily without any specialised methods. AAH03166 to AAH13628 and  
CC AAH13633 to AAH18742 represent human cDNA sequences; AAB92446 to  
CC AAB95893 represent human amino acid sequences; and AAH13629 to AAH13632  
CC represent oligonucleotides, all of which are used in the exemplification  
CC of the present invention.  
XX  
SQ Sequence 353 AA:  
  
Query Match 69.2%; Score 252; DB 22; Length 353;  
Best Local Similarity 99.7%; Pred. No. 3.7e-247;  
Matches 352; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
  
OY 12 MRLLSVLLGTPAPYVAVMGVRLSAPLRFYQALDDRLCYVYOSVLPFFFNNG 71  
DB 1 MYLLPSVVLGAPYVAVMGVRLSAPLRFYQALDDRLCYVYOSVLPFFFNNG 60  
  
OY 72 VOILLYGDLPKKNENIYLANHSTVDWIVADILAIKRONALGHVRYLKEGLKMLPLYGC 131  
DB 61 VQILLYGDLPKKNENIYLANHSTVDWIVADILAIKRONALGHVRYLKEGLKMLPLYGC 120  
  
OY 132 YFAGHGIVYKRSKAKFNEKEMRNKLOSYYDAGTPMLVTFPESTRNPEOTKYLASQAF 191  
DB 121 YFAGHGIVYKRSKAKFNEKEMRNKLOSYYDAGTPMLVTFPESTRNPEOTKYLASQAF 180  
  
OY 192 AAGRGIAVLKHLVTPRIKATHVAFDCMKNYLDAIYVTVVYEGKDDGGORRESPTMTFEL 251  
DB 181 AAGRGIAVLKHLVTPRIKATHVAFDCMKNYLDAIYVTVVYEGKDDGGORRESPTMTFEL 240  
  
OY 252 CKECPKIHIDIKDKDVEEEOEHMRRLHERFEIKDKMLIEFYSPDERRRKPPGKS 311  
DB 241 CKECPKIHIDIKDKDVEEEOEHMRRLHERFEIKDKMLIEFYSPDERRRKPPGKS 300  
  
OY 312 VNSKLSIKTLPSPMLILSGITAGMLMTDAGRKLYVNTWYIGLLGLCMTWTIKA 364  
DB 301 VNSKLSIKTLPSPMLILSGITAGMLMTDAGRKLYVNTWYIGLLGLCMTWTIKA 353  
  
RESULT 11  
ID AAY36744 standard; Protein; 291 AA.  
XX AAY36744;  
XX  
XX AAY36740;  
XX  
XX 27-SEP-1999 (first entry)  
XX  
XX PGI splice variant allele protein sequence.  
XX

KW PGI gene; diallelic marker; human; PSA; PGI-related diallelic marker;  
KM cancer; prostate cancer; diagnosis; therapy; prostate specific antigen.  
XX  
XX Homo sapiens.  
XX  
XX WO932644-A2.  
XX  
XX 01-JUL-1999.  
XX  
XX  
XX 22-DEC-1998; 98WO-1B02133.  
XX  
XX 09-SEP-1998; 98US-0099658.  
XX  
XX 22-DEC-1997; 97US-0996306.  
XX  
XX (GESP ) GENSET.  
XX  
XX Blumenfeld M, Bougueleret L, Chumakov I, Cohen D;  
PI WPI: 1999-405178/34.  
XX  
XX DR N-PsDB; AA00970.  
XX  
XX  
XX Use of a prostate cancer associated gene and diallelic markers  
PT derived from it  
XX  
XX  
XX Claim 7; Page 250-251; 385pp; English.  
XX  
XX The invention relates to a mammalian PGI gene and protein, and a set of  
CC PGI diallelic markers. The PGI polynucleotide and diallelic markers are  
CC used in a hybridisation assay, a sequencing assay, or in an  
CC allele-specific amplification assay for determining the identity of a  
CC nucleotide at a PGI-related diallelic marker. The methods can be used to  
CC detect and to assess the risk of developing cancer or prostate cancer.  
CC Early-stage diagnosis of prostate cancer relies on prostate specific  
CC antigen (PSA) dosage. However, the effectiveness of this is limited due  
CC to its inability to discriminate between malignant and non-malignant  
CC affections of the organ. A need exists for both a reliable diagnostic  
CC procedure which would enable early-stage diagnosis, and for preventative  
CC and curative treatments of the disease. The PGI gene can be used for  
CC detection of prostate cancer, and the risk of developing it in the  
CC future, and can also be used to determine therapies for the disease.  
XX  
SQ Sequence 291 AA:  
  
Query Match 63.2%; Score 230; DB 20; Length 291;  
Best Local Similarity 100.0%; Pred. No. 6.7e-225;  
Matches 230; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
OY 135 QHGGIVYKRSKAKFNEKEMRNKLOSYYDAGTPMLVTFPESTRNPEOTKYLASQAF 194  
DB 62 QHGGIVYKRSKAKFNEKEMRNKLOSYYDAGTPMLVTFPESTRNPEOTKYLASQAF 121  
  
OY 195 RGLAVLKHVLPRIKATHVAFDCMKNYLDAIYVTVVYEGKDDGGORRESPTMTFEL 254  
DB 122 RGLAVLKHVLPRIKATHVAFDCMKNYLDAIYVTVVYEGKDDGGORRESPTMTFEL 181  
  
OY 255 CKECPKIHIDIKDKDVEEEOEHMRRLHERFEIKDKMLIEFYSPDERRRKPPGKS 314  
DB 182 CKECPKIHIDIKDKDVEEEOEHMRRLHERFEIKDKMLIEFYSPDERRRKPPGKS 241  
  
OY 315 KLSIKTLPSPMLILSGITAGMLMTDAGRKLYVNTWYIGLLGLCMTWTIKA 364  
DB 242 KLSIKTLPSPMLILSGITAGMLMTDAGRKLYVNTWYIGLLGLCMTWTIKA 291  
  
RESULT 12  
ID AAY36740 standard; Protein; 228 AA.  
XX AAY36740;  
XX  
XX AAY36740;  
XX  
XX 27-SEP-1999 (first entry)  
XX  
XX

DE Truncated PGI protein sequence.  
XX  
KW PGI gene; biallelic marker; human; PSA; PGI-related biallelic marker;  
XX cancer; prostate cancer; diagnosis; therapy; prostate specific antigen.  
XX  
OS Homo sapiens.  
XX  
PN WO932644-A2.  
XX  
PD 01-JUL-1999.  
XX  
PE 22-DEC-1998; 98WO-IB02133.  
XX  
PR 09-SEP-1998; 98US-0099658.  
XX 22-DEC-1997; 97US-0996306.  
XX  
PA (GEST ) GENSET.  
XX  
PI Blumenfeld M, Bougueleret L, Chumakov I, Cohen D;  
XX  
DR WPI: 1999-405178/34.  
DR N-PSDB; AAZ00926.  
XX  
PT Use of a prostate cancer associated gene and biallelic markers  
XX derived from it  
XX  
PS Claim 7: Page 208; 385pp; English.  
XX  
XX The invention relates to a mammalian PGI gene and protein, and a set of  
CC PGI biallelic markers. The PGI polynucleotide and biallelic markers are  
CC used in a hybridisation assay, a sequencing assay, or in an  
CC allele-specific amplification assay for determining the identity of a  
CC nucleotide at a PGI-related biallelic marker. The methods can be used to  
CC detect and to assess the risk of developing cancer or prostate cancer.  
CC Early-stage diagnosis of prostate cancer relies on prostate specific  
CC antigen (PSA) dosage. However, the effectiveness of this is limited due  
CC to its inability to discriminate between malignant and non-malignant  
CC afflictions of the organ. A need exists for both a reliable diagnostic  
CC procedure which would enable early-stage diagnosis, and for preventative  
CC and curative treatments of the disease. The PGI gene can be used for  
CC detection of prostate cancer, and the risk of developing it in the  
CC future, and can also be used to determine therapies for the disease.  
XX  
XX Sequence 228 AA:  
SQ

Query Match 62.6%; Score 228; DB 20; Length 228;  
Best Local Similarity 100.0%; Pred. No. 5.7e-223;  
Matches 228; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 12 MYLLPSVLLGTAPTVLAWGVWRLSAFLPARFYQALDDRLXCVYQSMVLEFFENYTG 71  
Db 1 myllpsvllgtaptvylawgvwrlsafflparfyqalddrlxcvysmvleffenytg 60  
QY 72 VQILLKGDLPKKNENTITLNNHSTVDWIVADILATRONALGNRYVLAEGKMLPLKYC 131  
Db 61 vqillkgdlpkknentitlannhstvdwivadilatronalngrvylakgmlplkygc 120  
QY 132 YFQNHGIIYVRSKAFNEKEMRNKLOSIVDAGTPMYLVIPPEGRYRPEOTKVLASQAF 191  
Db 121 yfqnghiiyvrsakfnekemrnklosivdagtprmylvippegryrpeotkvlasaqaf 180  
QY 192 AAGRGIAVLKHLVLPRIKATHVAFDCMKNYLDAIYDVTYVVEGKDDG 239  
Db 181 aagrgiavlkhvlprikathvafdcmknyldaiydvtyvvegkddg 228

RESULT 13  
AA36745  
ID AAY36745 standard; Protein; 261 AA.  
XX  
AC AAY36745;  
XX

DT 27-SEP-1999 (first entry)  
XX  
DE PGI splice variant allele protein sequence.  
XX  
KW PGI gene; biallelic marker; human; PSA; PGI-related biallelic marker;  
XX cancer; prostate cancer; diagnosis; therapy; prostate specific antigen.  
XX  
OS Homo sapiens.  
XX  
PN WO932644-A2.  
XX  
PD 01-JUL-1999.  
XX  
PE 22-DEC-1998; 98WO-IB02133.  
XX  
PR 09-SEP-1998; 98US-0099658.  
XX 22-DEC-1997; 97US-0996306.  
XX  
PA (GEST ) GENSET.  
XX  
PI Blumenfeld M, Bougueleret L, Chumakov I, Cohen D;  
XX  
DR WPI: 1999-405178/34.  
DR N-PSDB; AAZ00971.  
XX  
PT Use of a prostate cancer associated gene and biallelic markers  
XX derived from it  
XX  
PS Claim 7: Page 251; 385pp; English.  
XX  
XX The invention relates to a mammalian PGI gene and protein, and a set of  
CC PGI biallelic markers. The PGI polynucleotide and biallelic markers are  
CC used in a hybridisation assay, a sequencing assay, or in an  
CC allele-specific amplification assay for determining the identity of a  
CC nucleotide at a PGI-related biallelic marker. The methods can be used to  
CC detect and to assess the risk of developing cancer or prostate cancer.  
CC Early-stage diagnosis of prostate cancer relies on prostate specific  
CC antigen (PSA) dosage. However, the effectiveness of this is limited due  
CC to its inability to discriminate between malignant and non-malignant  
CC afflictions of the organ. A need exists for both a reliable diagnostic  
CC procedure which would enable early-stage diagnosis, and for preventative  
CC and curative treatments of the disease. The PGI gene can be used for  
CC detection of prostate cancer, and the risk of developing it in the  
CC future, and can also be used to determine therapies for the disease.  
XX  
XX Sequence 261 AA:  
SQ

Query Match 54.7%; Score 199; DB 20; Length 261;  
Best Local Similarity 100.0%; Pred. No. 1.8e-193;  
Matches 199; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 166 MYLVIPPEGRYRPEOTKVLASQAFQAFQGLVNLKHLVLPRIKATHVAFDCMKNYLDAI 225  
Db 63 mylvippegryrpeotkvlasaqafaqglvnlkhvlprikathvafdcmknyldai 122  
QY 226 YDVTYVVEGKDDGQRRESPTMTTEFLCKECPKIHIDRIDKKVDPEQSHMRWLHERF 285  
Db 123 ydvtyvvegkddgqrresptmtteflckepkhihridriddkvdpeeqshmrwlherf 182  
QY 286 EIKDKMLIERESPDERRRKRFPGKSVNSKLSIKTLPSMLISGLTAGMLMDAGRKLY 345  
Db 183 eikdkmliefespderrkrfpgksvnskslksltlpsmlisgltagmlmdagrkly 242  
QY 346 VNTWIGTILGCLAWPTKA 364  
Db 243 vntwigtilgclawptka 261

RESULT 14  
AA36753  
ID AAY36753 standard; Protein; 185 AA.  
XX

AC AAY36753;  
XX  
XX 27-SEP-1999 (first entry)  
XX  
DE PGI splice variant allele protein sequence.  
XX  
XX PGI gene; biallelic marker; human; PSA; PGI-related biallelic marker;  
KW cancer; prostate cancer; diagnosis; therapy; prostate specific antigen.  
XX  
OS Homo sapiens.  
XX  
PN WO932644-A2.  
XX  
PD 01-JUL-1999.  
XX  
PF 22-DEC-1998; 98WO-IB02133.  
XX  
PR 09-SEP-1998; 98US-0099658.  
XX 22-DEC-1997; 97US-0996306.  
XX  
PA (GEST ) GENSET.  
XX  
PI Blumenfeld M, Bougueleret L, Chumakov I, Cohen D;  
XX WPI: 1999-405178/34.  
DR N-PSDB; AA200979.  
XX  
XX use of a prostate cancer associated gene and biallelic markers  
PT derived from it  
XX  
XX Claim 7; Page 255; 385pp; English.  
XX  
XX The invention relates to a mammalian PGI gene and protein, and a set of  
CC PGI biallelic markers. The PGI polynucleotide and biallelic markers are  
CC used in a hybridisation assay, a sequencing assay, or in an  
CC allele-specific amplification assay for determining the identity of a  
CC nucleotide at a PGI-related biallelic marker. The methods can be used to  
CC detect and to assess the risk of developing cancer or prostate cancer.  
CC Early-stage diagnosis of prostate cancer relies on prostate specific  
CC antigen (PSA) dosage. However, the effectiveness of this is limited due  
CC to its inability to discriminate between malignant and non-malignant  
CC affections of the organ. A need exists for both a reliable diagnostic  
CC procedure which would enable early-stage diagnosis, and for preventative  
CC and curative treatments of the disease. The PGI gene can be used for  
CC detection of prostate cancer, and the risk of developing it in the  
CC future, and can also be used to determine therapies for the disease.  
XX  
XX Sequence 185 AA;  
SQ

Query Match 50.8%; Score 185; DB 20; Length 185;  
Best Local Similarity 100.0%; Pred. No. 2.1e-179;  
Matches 185; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
OY 12 MRYLLPSVVLGTAPRYVLAAGWRLLSAFLPARFYQALDDRLXCYQSWLFEFFENYTG 71  
DB 1 mryllpsvvlgtapryvlawgwrlsalfparfyqalddrlxcyqswvlffenytg 60  
OY 72 VOILLYGDLPKKNENITTYLANHOSVDWIVADILAIRONALGHVRYVLEKGLKWLPLYGC 131  
DB 61 vqilllygdlpkknenitlylanhbstvdwivadilairgnalghvryvlekgllkwlplygc 120  
OY 132 YFAOHGIIYKRSKAKNEKEMRNKLSQSYDAGTGMTLVIFPECTRYNPEOTKVLASQA 191  
DB 121 yfaohggiykrskaknekeemrnlqsyvdagtpmylvifpegttrynpeotkvlasaqaf 180  
OY 192 AAORG 196  
DB 181 aadrg 185

RESULT 15  
AAY36751

ID AAY36751 standard; Protein; 315 AA.  
XX  
XX AAY36751;  
XX  
XX 27-SEP-1999 (first entry)  
XX  
DE PGI splice variant allele protein sequence.  
XX  
XX PGI gene; biallelic marker; human; PSA; PGI-related biallelic marker;  
KW cancer; prostate cancer; diagnosis; therapy; prostate specific antigen.  
XX  
XX Homo sapiens.  
XX  
PN WO932644-A2.  
XX  
PD 01-JUL-1999.  
XX  
PF 22-DEC-1998; 98WO-IB02133.  
XX  
PR 09-SEP-1998; 98US-0099658.  
XX 22-DEC-1997; 97US-0996306.  
XX  
PA (GEST ) GENSET.  
XX  
PI Blumenfeld M, Bougueleret L, Chumakov I, Cohen D;  
XX WPI: 1999-405178/34.  
DR N-PSDB; AA200977.  
XX  
XX use of a prostate cancer associated gene and biallelic markers  
PT derived from it  
XX  
XX Claim 7; Page 253-254; 385pp; English.  
XX  
XX The invention relates to a mammalian PGI gene and protein, and a set of  
CC PGI biallelic markers. The PGI polynucleotide and biallelic markers are  
CC used in a hybridisation assay, a sequencing assay, or in an  
CC allele-specific amplification assay for determining the identity of a  
CC nucleotide at a PGI-related biallelic marker. The methods can be used to  
CC detect and to assess the risk of developing cancer or prostate cancer.  
CC Early-stage diagnosis of prostate cancer relies on prostate specific  
CC antigen (PSA) dosage. However, the effectiveness of this is limited due  
CC to its inability to discriminate between malignant and non-malignant  
CC affections of the organ. A need exists for both a reliable diagnostic  
CC procedure which would enable early-stage diagnosis, and for preventative  
CC and curative treatments of the disease. The PGI gene can be used for  
CC detection of prostate cancer, and the risk of developing it in the  
CC future, and can also be used to determine therapies for the disease.  
XX  
XX Sequence 315 AA;  
SQ

Query Match 50.8%; Score 185; DB 20; Length 315;  
Best Local Similarity 100.0%; Pred. No. 3.5e-179;  
Matches 185; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
OY 12 MRYLLPSVVLGTAPRYVLAAGWRLLSAFLPARFYQALDDRLXCYQSWLFEFFENYTG 71  
DB 1 mryllpsvvlgtapryvlawgwrlsalfparfyqalddrlxcyqswvlffenytg 60  
OY 72 VOILLYGDLPKKNENITTYLANHOSVDWIVADILAIRONALGHVRYVLEKGLKWLPLYGC 131  
DB 61 vqilllygdlpkknenitlylanhbstvdwivadilairgnalghvryvlekgllkwlplygc 120  
OY 132 YFAOHGIIYKRSKAKNEKEMRNKLSQSYDAGTGMTLVIFPECTRYNPEOTKVLASQA 191  
DB 121 yfaohggiykrskaknekeemrnlqsyvdagtpmylvifpegttrynpeotkvlasaqaf 180  
OY 192 AAORG 196  
DB 181 aadrg 185

Wed Aug 28 13:17:04 2002

us-09-853-526-5.rag

Page 12

Search completed: August 28, 2002, 11:18:44  
Job time: 582 sec

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 28, 2002, 11:19:21 ; Search time 30.82 Seconds

(without alignments)  
288.479 Million cell updates/sec

Title: US-09-853-526-5

Perfect score: 364

Sequence: 1 MLSTLVHRYSMRYLLPSVY.....YVNTWYIGTLGLWYIKR 364

Scoring table: OLIGO

Searched: 231628 seqs, 24425594 residues

Word size : 0

Total number of hits satisfying chosen parameters: 231628

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : Issued\_Patents\_AA.\*

1: /cgn2\_6/ptodata/2/1aa/5A\_COMB.pep.\*  
2: /cgn2\_6/ptodata/2/1aa/5B\_COMB.pep.\*  
3: /cgn2\_6/ptodata/2/1aa/6A\_COMB.pep.\*  
4: /cgn2\_6/ptodata/2/1aa/6B\_COMB.pep.\*  
5: /cgn2\_6/ptodata/2/1aa/PCtUS\_COMB.pep.\*  
6: /cgn2\_6/ptodata/2/1aa/Backfillseq1.pep.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	364	100.0	364	2	US-08-996-306-5
2	364	100.0	364	4	US-09-338-907-5
3	364	100.0	364	4	US-09-218-207-5
4	353	97.0	353	2	US-08-996-306-4
5	353	97.0	353	4	US-09-338-907-4
6	353	97.0	353	4	US-09-218-207-4
7	230	63.2	291	4	US-09-338-907-127
8	230	63.2	291	4	US-09-218-207-127
9	228	62.6	228	4	US-09-338-907-70
10	228	62.6	228	4	US-09-218-207-70
11	199	54.7	261	4	US-09-338-907-128
12	199	54.7	261	4	US-09-218-207-128
13	185	50.8	185	4	US-09-338-907-136
14	185	50.8	185	4	US-09-218-207-136
15	185	50.8	185	4	US-09-338-907-134
16	185	50.8	185	4	US-09-338-907-134
17	184	50.5	300	4	US-09-218-207-135
18	184	50.5	300	4	US-09-338-907-135
19	154	42.3	182	4	US-09-218-207-133
20	154	42.3	182	4	US-09-338-907-133
21	116	31.9	238	4	US-09-218-207-126
22	116	31.9	238	4	US-09-338-907-126
23	85	23.4	97	4	US-09-218-207-132
24	85	23.4	97	4	US-09-338-907-132
25	62	17.0	66	4	US-09-218-207-131
26	62	17.0	66	4	US-09-338-907-131
27	62	17.0	68	4	US-09-338-907-130

28	62	17.0	68	4	US-09-218-207-130	Sequence 130, App
29	62	17.0	77	4	US-09-338-907-125	Sequence 125, App
30	62	17.0	77	4	US-09-218-207-125	Sequence 125, App
31	62	17.0	90	4	US-09-338-907-129	Sequence 129, App
32	62	17.0	90	4	US-09-218-207-129	Sequence 129, App
33	32	8.8	354	4	US-09-338-907-74	Sequence 74, App1
34	32	8.8	354	4	US-09-218-207-74	Sequence 74, App1
35	9	2.5	9	2	US-08-996-306-15	Sequence 15, App1
36	9	2.5	9	4	US-09-338-907-15	Sequence 15, App1
37	9	2.5	9	4	US-09-218-207-15	Sequence 15, App1
38	8	2.2	9	2	US-08-996-306-18	Sequence 18, App1
39	8	2.2	9	4	US-09-338-907-18	Sequence 18, App1
40	8	2.2	9	4	US-09-218-207-18	Sequence 18, App1
41	7	1.9	12	1	US-08-224-625-1	Sequence 1, App1
42	7	1.9	12	2	US-08-254-404-1	Sequence 1, App1
43	7	1.9	12	2	US-08-327-451E-1	Sequence 1, App1
44	7	1.9	12	2	US-08-458-109-1	Sequence 1, App1
45	7	1.9	12	3	US-08-231-196-1	Sequence 1, App1

#### ALIGNMENTS

RESULT 1  
US-08-996-306-5  
; Sequence 5, Application US/08996306  
; Patent No. 5945522  
; GENERAL INFORMATION:  
; APPLICANT: Cohen, Daniel  
; APPLICANT: Chumakov, Ilya  
; APPLICANT: Blumenfeld, Martia  
; APPLICANT: Bougueleret, Lydie  
; TITLE OF INVENTION: Prostate cancer gene  
; NUMBER OF SEQUENCES: 68  
; CORRESPONDENCE ADDRESSES:  
; ADDRESSEE: Knobb, Martens, Olson & Bear  
; STREET: 501 West Broadway  
; CITY: San Diego  
; STATE: California  
; COUNTRY: USA  
; ZIP: 92101-3505  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy Disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: Win95  
; SOFTWARE: Word  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/996,306  
; FILING DATE:  
; CLASSIFICATION:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Israel, Ned A.  
; REGISTRATION NUMBER: 29,655  
; REFERENCE/DOCKET NUMBER: GENSET, 018A  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (619) 235-8550  
; TELEFAX: (619) 235-0176  
; INFORMATION FOR SEQ ID NO: 5:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 364 amino acids  
; TYPE: AMINO ACID  
; STRANDEDNESS: SINGLE  
; TOPOLOGY: LINEAR  
; MOLECULE TYPE: PROTEIN  
; ORIGINAL SOURCE:  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: potential protein  
; LOCATION: 1..364  
; US-08-996-306-5

Query Match 100.0%; Score 364; DB 2; Length 364;

```
Best Local Similarity 100.0%; Pred. No. 0;
Matches 364; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 MLSTLVHHTYSMRYYLLPSVLLGTAFTYVLAAGWWRLLSAFLPARFYQALDRLCYVOS 60
    |||
Db 1 MLSTLVHHTYSMRYYLLPSVLLGTAFTYVLAAGWWRLLSAFLPARFYQALDRLCYVOS 60

Oy 61 MYLFFENYTGVOILLYGDLPRKNKENIITLANHGSTVDMIVADILAIRONALGHVRYVAK 120
    |||
Db 61 MYLFFENYTGVOILLYGDLPRKNKENIITLANHGSTVDMIVADILAIRONALGHVRYVAK 120

Oy 121 EGIKMLPLYGCTFAOHGGIYVKRSKAFNEKEMRNKLQSYVDAGTPMYLVIFPEGTRYNPE 180
    |||
Db 121 EGIKMLPLYGCTFAOHGGIYVKRSKAFNEKEMRNKLQSYVDAGTPMYLVIFPEGTRYNPE 180

Oy 181 QTKVLSASQAFPAOGRALAVLKHVLPRIKATHVAFDCMKNYDAIYDVYVYEGKDDGQ 240
    |||
Db 181 QTKVLSASQAFPAOGRALAVLKHVLPRIKATHVAFDCMKNYDAIYDVYVYEGKDDGQ 240

Oy 241 RRESPTMEFLCKECPKIHIDRIDKDVPEBOEHMRMLHERFEIKDKMLIEFYESP 300
    |||
Db 241 RRESPTMEFLCKECPKIHIDRIDKDVPEBOEHMRMLHERFEIKDKMLIEFYESP 300

Oy 301 PERRRKPEKSVNSKLSIKKTLPSMLISGLTAGMLMTDAGRKLVTYNTWYIGTLGCLMV 360
    |||
Db 301 PERRRKPEKSVNSKLSIKKTLPSMLISGLTAGMLMTDAGRKLVTYNTWYIGTLGCLMV 360

Oy 361 TIRKA 364
    |||
Db 361 TIRKA 364

RESULT 2
US-09-338-907-5
; Sequence 5, Application US/09338907
; Patent No. 6265546
; GENERAL INFORMATION:
; APPLICANT: Cohen, Daniel
; APPLICANT: Blumenfeld, Marta
; APPLICANT: Ilya, Chumakov
; APPLICANT: Bougueleret, Lydie
; TITLE OF INVENTION: PROSTATE CANCER GENE
; FILE REFERENCE: GENSET 18CPICP
; CURRENT APPLICATION NUMBER: US/09/338,907
; CURRENT FILING DATE: 1999-06-23
; EARLIER APPLICATION NUMBER: 08/996,306
; EARLIER FILING DATE: 1997-12-22
; EARLIER APPLICATION NUMBER: 60/099,658
; EARLIER FILING DATE: 1998-09-09
; EARLIER APPLICATION NUMBER: 09/218,207
; EARLIER FILING DATE: 1998-12-22
; NUMBER OF SEQ ID NOS: 578
; SOFTWARE: Patent.pm
; SEQ ID NO 5
; LENGTH: 364
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-338-907-5

Query Match 100.0%; Score 364; DB 4; Length 364;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 364; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 MLSTLVHHTYSMRYYLLPSVLLGTAFTYVLAAGWWRLLSAFLPARFYQALDRLCYVOS 60
    |||
Db 1 MLSTLVHHTYSMRYYLLPSVLLGTAFTYVLAAGWWRLLSAFLPARFYQALDRLCYVOS 60

Oy 61 MYLFFENYTGVOILLYGDLPRKNKENIITLANHGSTVDMIVADILAIRONALGHVRYVAK 120
    |||
Db 61 MYLFFENYTGVOILLYGDLPRKNKENIITLANHGSTVDMIVADILAIRONALGHVRYVAK 120

Oy 121 EGIKMLPLYGCTFAOHGGIYVKRSKAFNEKEMRNKLQSYVDAGTPMYLVIFPEGTRYNPE 180
    |||
Db 121 EGIKMLPLYGCTFAOHGGIYVKRSKAFNEKEMRNKLQSYVDAGTPMYLVIFPEGTRYNPE 180
```

```

    |||
Db 121 EGIKMLPLYGCTFAOHGGIYVKRSKAFNEKEMRNKLQSYVDAGTPMYLVIFPEGTRYNPE 180

Oy 181 QTKVLSASQAFPAOGRALAVLKHVLPRIKATHVAFDCMKNYDAIYDVYVYEGKDDGQ 240
    |||
Db 181 QTKVLSASQAFPAOGRALAVLKHVLPRIKATHVAFDCMKNYDAIYDVYVYEGKDDGQ 240

Oy 241 RRESPTMEFLCKECPKIHIDRIDKDVPEBOEHMRMLHERFEIKDKMLIEFYESP 300
    |||
Db 241 RRESPTMEFLCKECPKIHIDRIDKDVPEBOEHMRMLHERFEIKDKMLIEFYESP 300

Oy 301 PERRRKPEKSVNSKLSIKKTLPSMLISGLTAGMLMTDAGRKLVTYNTWYIGTLGCLMV 360
    |||
Db 301 PERRRKPEKSVNSKLSIKKTLPSMLISGLTAGMLMTDAGRKLVTYNTWYIGTLGCLMV 360

Oy 361 TIRKA 364
    |||
Db 361 TIRKA 364

RESULT 3
US-09-218-207-5
; Sequence 5, Application US/09218207
; Patent No. 6346381
; GENERAL INFORMATION:
; APPLICANT: Cohen, Daniel
; APPLICANT: Blumenfeld, Marta
; APPLICANT: Ilya, Chumakov
; APPLICANT: Bougueleret, Lydie
; TITLE OF INVENTION: Prostate cancer gene
; FILE REFERENCE: GENSET 018CP1
; CURRENT APPLICATION NUMBER: US/09/218,207
; CURRENT FILING DATE: 1998-12-22
; EARLIER APPLICATION NUMBER: 08/996,306
; EARLIER FILING DATE: 1997-12-22
; EARLIER APPLICATION NUMBER: 60/099,658
; EARLIER FILING DATE: 1998-09-09
; NUMBER OF SEQ ID NOS: 578
; SOFTWARE: Patent.pm
; SEQ ID NO 5
; LENGTH: 364
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-218-207-5

Query Match 100.0%; Score 364; DB 4; Length 364;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 364; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 MLSTLVHHTYSMRYYLLPSVLLGTAFTYVLAAGWWRLLSAFLPARFYQALDRLCYVOS 60
    |||
Db 1 MLSTLVHHTYSMRYYLLPSVLLGTAFTYVLAAGWWRLLSAFLPARFYQALDRLCYVOS 60

Oy 61 MYLFFENYTGVOILLYGDLPRKNKENIITLANHGSTVDMIVADILAIRONALGHVRYVAK 120
    |||
Db 61 MYLFFENYTGVOILLYGDLPRKNKENIITLANHGSTVDMIVADILAIRONALGHVRYVAK 120

Oy 121 EGIKMLPLYGCTFAOHGGIYVKRSKAFNEKEMRNKLQSYVDAGTPMYLVIFPEGTRYNPE 180
    |||
Db 121 EGIKMLPLYGCTFAOHGGIYVKRSKAFNEKEMRNKLQSYVDAGTPMYLVIFPEGTRYNPE 180

Oy 181 QTKVLSASQAFPAOGRALAVLKHVLPRIKATHVAFDCMKNYDAIYDVYVYEGKDDGQ 240
    |||
Db 181 QTKVLSASQAFPAOGRALAVLKHVLPRIKATHVAFDCMKNYDAIYDVYVYEGKDDGQ 240

Oy 241 RRESPTMEFLCKECPKIHIDRIDKDVPEBOEHMRMLHERFEIKDKMLIEFYESP 300
    |||
Db 241 RRESPTMEFLCKECPKIHIDRIDKDVPEBOEHMRMLHERFEIKDKMLIEFYESP 300

Oy 301 PERRRKPEKSVNSKLSIKKTLPSMLISGLTAGMLMTDAGRKLVTYNTWYIGTLGCLMV 360
    |||
Db 301 PERRRKPEKSVNSKLSIKKTLPSMLISGLTAGMLMTDAGRKLVTYNTWYIGTLGCLMV 360
```

OY 361 TIKa 364  
DB 361 TIKa 364

RESULT 4  
US-08-996-306-4  
Sequence 4, Application US/08996306  
Patent No. 5945522  
GENERAL INFORMATION:  
APPLICANT: Cohen, Daniel  
APPLICANT: Chumakov, Ilya  
APPLICANT: Blumenfeld, Marta  
APPLICANT: Bouguenoret, Lydie  
TITLE OF INVENTION: Prostate cancer gene  
NUMBER OF SEQUENCES: 68  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Knobbe, Martens, Olson & Bear  
STREET: 501 West Broadway  
CITY: San Diego  
STATE: California  
COUNTRY: USA  
ZIP: 92101-3505  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy Disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: Win95  
SOFTWARE: Word  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/996,306  
FILING DATE:  
CLASSIFICATION:  
ATTORNEY/AGENT INFORMATION:  
NAME: Israelien, Ned A.  
REGISTRATION NUMBER: 29,655  
REFERENCE/DOCKET NUMBER: GENSET.018A  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (619) 235-8550  
TELEFAX: (619) 235-0176  
INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 353 amino acids  
TYPE: AMINO ACID  
STRANDEDNESS: SINGLE  
TOPOLOGY: LINEAR  
MOLECULE TYPE: PROTEIN  
ORIGINAL SOURCE:  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: potential Transmembrane helix  
LOCATION: 1..33  
IDENTIFICATION METHOD: Rao and Argos method  
FEATURE:  
NAME/KEY: potential Transmembrane helix  
LOCATION: 4..20  
IDENTIFICATION METHOD: Klein, Kanehisa and Delisi method  
FEATURE:  
NAME/KEY: potential Transmembrane helix  
LOCATION: 4..24  
IDENTIFICATION METHOD: Eisenberg, Schwarz, Komaromy  
IDENTIFICATION METHOD: and Wall method  
FEATURE:  
NAME/KEY: potential N-myristoylation site  
LOCATION: 12  
IDENTIFICATION METHOD: prosite match  
FEATURE:  
NAME/KEY: potential Transmembrane helix  
LOCATION: 50..70  
IDENTIFICATION METHOD: Eisenberg, Schwarz, Komaromy and Wall method  
FEATURE:  
NAME/KEY: potential N-glycosylation site  
LOCATION: 57

IDENTIFICATION METHOD: prosite match  
FEATURE:  
NAME/KEY: potential Transmembrane helix  
LOCATION: 76..96  
IDENTIFICATION METHOD: Eisenberg, Schwarz, Komaromy and Wall method  
FEATURE:  
NAME/KEY: potential Tyrosine kinase phosphorylation site  
LOCATION: 78  
IDENTIFICATION METHOD: prosite match  
FEATURE:  
NAME/KEY: potential Casein kinase II phosphorylation site  
LOCATION: 84  
IDENTIFICATION METHOD: prosite match  
FEATURE:  
NAME/KEY: potential Leucine zipper pattern  
LOCATION: 94..115  
IDENTIFICATION METHOD: prosite match  
FEATURE:  
NAME/KEY: potential N-myristoylation site  
LOCATION: 119  
IDENTIFICATION METHOD: prosite match  
FEATURE:  
NAME/KEY: potential Protein kinase C phosphorylation site  
LOCATION: 133  
IDENTIFICATION METHOD: prosite match  
FEATURE:  
NAME/KEY: potential Casein kinase II phosphorylation site  
LOCATION: 147  
IDENTIFICATION METHOD: prosite match  
FEATURE:  
NAME/KEY: potential Protein kinase C phosphorylation site  
LOCATION: 194  
IDENTIFICATION METHOD: prosite match  
FEATURE:  
NAME/KEY: potential Tyrosine kinase phosphorylation site  
LOCATION: 215  
IDENTIFICATION METHOD: prosite match  
FEATURE:  
NAME/KEY: potential Tyrosine sulfatation site  
LOCATION: 221  
IDENTIFICATION METHOD: prosite match  
FEATURE:  
NAME/KEY: potential CAMP- and cGMP-dependent protein kinase phosphorylation  
LOCATION: 233  
IDENTIFICATION METHOD: prosite match  
FEATURE:  
NAME/KEY: potential Casein kinase II phosphorylation site  
LOCATION: 235  
IDENTIFICATION METHOD: prosite match  
FEATURE:  
NAME/KEY: potential Protein kinase C phosphorylation site  
LOCATION: 306  
IDENTIFICATION METHOD: prosite match  
FEATURE:  
NAME/KEY: potential Transmembrane helix  
LOCATION: 310..330  
IDENTIFICATION METHOD: Eisenberg, Schwarz, Komaromy and Wall method  
FEATURE:  
NAME/KEY: potential N-myristoylation site  
LOCATION: 319  
IDENTIFICATION METHOD: prosite match  
FEATURE:  
NAME/KEY: potential N-myristoylation site  
LOCATION: 323  
IDENTIFICATION METHOD: prosite match  
FEATURE:  
NAME/KEY: potential Amidation site  
LOCATION: 329  
IDENTIFICATION METHOD: prosite match  
FEATURE:  
NAME/KEY: potential Transmembrane helix  
LOCATION: 333..353  
IDENTIFICATION METHOD: Eisenberg, Schwarz, Komaromy and Wall method

```

;
; FEATURE:
; NAME/KEY: potential N-myristoylation site
; LOCATION: 341
; IDENTIFICATION METHOD: prosite match
; FEATURE:
; NAME/KEY: potential protein kinase C phosphorylation site
; LOCATION: 350
; IDENTIFICATION METHOD: prosite match
; US-08-996-306-4

Query Match      97.0%: Score 353; DB 2; Length 353;
Best Local Similarity 100.0%: Pred. No. 0;
Matches 353; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 12 MRYLLPSVVLGTAFTYVLMGVRLSAFLPARFYQALDRLCYVQSNVLEFFENYTG 71
    |||||||
DB 1 MRYLLPSVVLGTAFTYVLMGVRLSAFLPARFYQALDRLCYVQSNVLEFFENYTG 60
    |||||||
QY 72 VOILLGDLPRKKNENIYYLANHOSYVDWIVADILAIKQNALGHVRYVLAEGKWLPLYGC 131
    |||||||
DB 61 VOILLGDLPRKKNENIYYLANHOSYVDWIVADILAIKQNALGHVRYVLAEGKWLPLYGC 120
    |||||||
QY 132 YFQHGSIYKRSKKEKEMRNKLSYVDAGTPMYLVIFPEGRVNPEDTKVLASQAF 191
    |||||||
DB 121 YFQHGSIYKRSKKEKEMRNKLSYVDAGTPMYLVIFPEGRVNPEDTKVLASQAF 180
    |||||||
QY 192 AARGGLAVLKHVLPRIKATVAFDCMKNYLDAIYDVTYVEGKDGQRRSEPTMTPEFL 251
    |||||||
DB 181 AARGGLAVLKHVLPRIKATVAFDCMKNYLDAIYDVTYVEGKDGQRRSEPTMTPEFL 240
    |||||||
QY 252 CKKCPKTHIHIDRIKDVDEEDHMRWLHERPEIKDKMLIEFYSPDPERKRRFPGKS 311
    |||||||
DB 241 CKKCPKTHIHIDRIKDVDEEDHMRWLHERPEIKDKMLIEFYSPDPERKRRFPGKS 300
    |||||||
QY 312 VNSKLSIKKTLPSMLISGLTAGMLMDAGRKLYVNTWITGLGCLWPTIKA 364
    |||||||
DB 301 VNSKLSIKKTLPSMLISGLTAGMLMDAGRKLYVNTWITGLGCLWPTIKA 353
    |||||||

RESULT 5
US-09-338-907-4
; Sequence 4, Application US/09338907
; Patent No. 6265546
; GENERAL INFORMATION:
; APPLICANT: Cohen, Daniel
; APPLICANT: Blumenfeld, Marla
; APPLICANT: Ilya, Chumakov
; APPLICANT: Bougueleret, Lydie
; TITLE OF INVENTION: PROSTATE CANCER GENE
; FILE REFERENCE: GENSET.18CPICP
; CURRENT APPLICATION NUMBER: US/09/338,907
; EARLIER FILING DATE: 1999-06-23
; EARLIER APPLICATION NUMBER: 08/996,306
; EARLIER FILING DATE: 1997-12-22
; EARLIER APPLICATION NUMBER: 60/099,658
; EARLIER FILING DATE: 1998-09-09
; EARLIER APPLICATION NUMBER: 09/218,207
; EARLIER FILING DATE: 1998-12-22
; NUMBER OF SEQ ID NOS: 578
; SOFTWARE: Patent.pm
; SEQ ID NO 4
; LENGTH: 353
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: HELIX
; LOCATION: 1..33
; OTHER INFORMATION: Rao and Argos identification method, potential helix
; FEATURE:
; NAME/KEY: HELIX
; LOCATION: 4..20
; OTHER INFORMATION: Klein, Kanehisa and Delist identification method, potential]

;
; FEATURE:
; NAME/KEY: HELIX
; LOCATION: 4..24
; OTHER INFORMATION: Eisenberg, Schwarz, Komarony, Wall identification method,
; FEATURE:
; NAME/KEY: MYRISTATE
; LOCATION: 12..16
; OTHER INFORMATION: prosite match
; FEATURE:
; NAME/KEY: HELIX
; LOCATION: 50..70
; OTHER INFORMATION: Eisenberg, Schwarz, Komarony, Wall identification method,
; FEATURE:
; NAME/KEY: CARBOHYD
; LOCATION: 57..59
; OTHER INFORMATION: prosite match
; FEATURE:
; NAME/KEY: HELIX
; LOCATION: 76..96
; OTHER INFORMATION: Eisenberg, Schwarz, Komarony, Wall identification method,
; FEATURE:
; NAME/KEY: PHOSPHORYLATION
; LOCATION: 78
; OTHER INFORMATION: potential tyrosine kinase site, prosite match
; FEATURE:
; NAME/KEY: PHOSPHORYLATION
; LOCATION: 84
; OTHER INFORMATION: potential caseline kinase II site, prosite match
; FEATURE:
; NAME/KEY: SITE
; LOCATION: 94..115
; OTHER INFORMATION: potential leucine zipper site, prosite match
; FEATURE:
; NAME/KEY: MYRISTATE
; LOCATION: 119..123
; OTHER INFORMATION: potential site, prosite match
; FEATURE:
; NAME/KEY: PHOSPHORYLATION
; LOCATION: 133
; OTHER INFORMATION: potential protein kinase C, prosite match
; FEATURE:
; NAME/KEY: PHOSPHORYLATION
; LOCATION: 147
; OTHER INFORMATION: potential caseline kinase II site, prosite match
; FEATURE:
; NAME/KEY: PHOSPHORYLATION
; LOCATION: 194
; OTHER INFORMATION: potential protein kinase C, prosite match
; FEATURE:
; NAME/KEY: PHOSPHORYLATION
; LOCATION: 215
; OTHER INFORMATION: potential tyrosine kinase site, prosite match
; FEATURE:
; NAME/KEY: SULFATATION
; LOCATION: 221
; OTHER INFORMATION: prosite match
; FEATURE:
; NAME/KEY: PHOSPHORYLATION
; LOCATION: 235
; OTHER INFORMATION: potential caseline kinase II site, prosite match
; FEATURE:
; NAME/KEY: PHOSPHORYLATION
; LOCATION: 306
; OTHER INFORMATION: potential protein kinase C, prosite match
; FEATURE:
; NAME/KEY: HELIX
; LOCATION: 310..330
; OTHER INFORMATION: Eisenberg, Schwarz, Komarony, Wall identification method,
; FEATURE:
```



```

? NAME/KEY: MYRISTATE
? LOCATION: 319..323
? OTHER INFORMATION: Prosite match
? FEATURE:
? NAME/KEY: MYRISTATE
? LOCATION: 323..327
? OTHER INFORMATION: Prosite match
? FEATURE:
? NAME/KEY: AMIDATION
? LOCATION: 329
? OTHER INFORMATION: Prosite match
? FEATURE:
? NAME/KEY: HELIX
? LOCATION: 333..353
? OTHER INFORMATION: Eisenberg, Schwarz, Komaromy, Wall identification method, F
? FEATURE:
? NAME/KEY: MYRISTATE
? LOCATION: 341..345
? OTHER INFORMATION: Prosite match
? FEATURE:
? NAME/KEY: PHOSPHORYLATION
? LOCATION: 350
? OTHER INFORMATION: potential protein kinase C, Prosite match
? OS-09-338-907-4

```

Query Match	97.0%;	Score 353;	DB 4;	Length 353;
Best Local Similarity	100.0%;	Pred. No. 0;		
Matches 353;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;

QY	12	MRYLPSVLLGTPATPYVLAMGVRRLLSAPFAEYQALDDRLCYQSWLPEFENYTG	71
Db	1	MRYLLPSVLLGTPATPYVLAMGVRRLLSAPFAEYQALDDRLCYQSWLPEFENYTG	60
QY	72	VOILLGDLPEKKNENITYLANHOSVDVMVADLLAIROMALGHVRYVLKELKMLPLYG	131
Db	61	VOILLGDLPEKKNENITYLANHOSVDVMVADLLAIROMALGHVRYVLKELKMLPLYG	120
QY	132	YFAOHGGIYVRSKAFENKEKRNKLOSYVADGPMYLYIPEEGRYRYPEOTKVLASQAF	191
Db	121	YFAOHGGIYVRSKAFENKEKRNKLOSYVADGPMYLYIPEEGRYRYPEOTKVLASQAF	180
QY	192	AAOGLAVLKHVLTPRIKATHVAFDCKKNYLDAIYDVTVVYEGKDDGGORRESEPTMEFL	251
Db	181	AAOGLAVLKHVLTPRIKATHVAFDCKKNYLDAIYDVTVVYEGKDDGGORRESEPTMEFL	240
QY	252	CKECPKLIHIDRLDOKKDVEPEOEHMRMYLHEREIYIDKMLIEFYESPDPERRRRPFGKS	311
Db	241	CKECPKLIHIDRLDOKKDVEPEOEHMRMYLHEREIYIDKMLIEFYESPDPERRRRPFGKS	300
QY	312	VNSKLSIIRKTLPSMLILSGLTAGMLMDAGRKLYVNTWVYTGTLGLCLMTVITKA	364
Db	301	VNSKLSIIRKTLPSMLILSGLTAGMLMDAGRKLYVNTWVYTGTLGLCLMTVITKA	353

```

RESULT 6
US-09-218-207-4
: Sequence 4, Application us/09218207
: Patent No. 6346381
GENERAL INFORMATION:
APPLICANT: Cohen, Daniel
APPLICANT: Blumenfeld, Marta
APPLICANT: Ilyu, Chumakov
APPLICANT: Bouquellet, Lydie
TITLE OF INVENTION: Prostate cancer gene
FILE REFERENCE: GENSET_018CP1
CURRENT APPLICATION NUMBER: us/09/218,207
CURRENT FILING DATE: 1998-12-22
EARLIER APPLICATION NUMBER: 08/9956,306
EARLIER FILING DATE: 1997-12-22
EARLIER APPLICATION NUMBER: 60/099,658
EARLIER FILING DATE: 1998-09-09
NUMBER OF SEQ ID NOS: 578

```

SOFTWARE: Patent.pm	SEQ ID NO 4	LENGTH: 353	TYPE: PRT	ORGANISM: Homo sapiens	FEATURE:	NAME/KEY: HELIX	LOCATION: 1..33	OTHER INFORMATION: Rao and Argos identification method, potential helix
					FEATURE:	NAME/KEY: HELIX	LOCATION: 4..20	OTHER INFORMATION: Klein, Kanehisa and Delisi identification method, potent
					FEATURE:	NAME/KEY: HELIX	LOCATION: 4..24	OTHER INFORMATION: Eisenberg, Schwarz, Komarony, wall identification method,
					FEATURE:	NAME/KEY: MYRISTATE	LOCATION: 12..16	OTHER INFORMATION: Prosite match
					FEATURE:	NAME/KEY: HELIX	LOCATION: 50..70	OTHER INFORMATION: Eisenberg, Schwarz, Komarony, wall identification method,
					FEATURE:	NAME/KEY: CARBOHYD	LOCATION: 57..59	OTHER INFORMATION: Prosite match
					FEATURE:	NAME/KEY: HELIX	LOCATION: 76..96	OTHER INFORMATION: Eisenberg, Schwarz, Komarony, wall identification method,
					FEATURE:	NAME/KEY: PHOSPHORYLATION	LOCATION: 78	OTHER INFORMATION: potential Tyrosine kinase site, Prosite match
					FEATURE:	NAME/KEY: SITE	LOCATION: 94..115	OTHER INFORMATION: potential Leucine zipper site, Prosite match
					FEATURE:	NAME/KEY: MYRISTATE	LOCATION: 119..123	OTHER INFORMATION: potential site, Prosite match
					FEATURE:	NAME/KEY: PHOSPHORYLATION	LOCATION: 133	OTHER INFORMATION: potential protein kinase C, Prosite match
					FEATURE:	NAME/KEY: PHOSPHORYLATION	LOCATION: 147	OTHER INFORMATION: potential caseine kinase II site, Prosite match
					FEATURE:	NAME/KEY: PHOSPHORYLATION	LOCATION: 194	OTHER INFORMATION: potential protein kinase C, Prosite match
					FEATURE:	NAME/KEY: PHOSPHORYLATION	LOCATION: 215	OTHER INFORMATION: potential Tyrosine kinase site, Prosite match
					FEATURE:	NAME/KEY: SULFATATION	LOCATION: 221	OTHER INFORMATION: Prosite match
					FEATURE:	NAME/KEY: PHOSPHORYLATION	LOCATION: 233	OTHER INFORMATION: potential cAMP and cGMP dependant protein kinase site,

```

; FEATURE:
; NAME/KEY: PHOSPHORYLATION
; LOCATION: 235
; OTHER INFORMATION: potential caseine kinase II site, Prosite match
; FEATURE:
; NAME/KEY: PHOSPHORYLATION
; LOCATION: 306
; OTHER INFORMATION: potential protein kinase C, Prosite match
; FEATURE:
; NAME/KEY: HELIX
; LOCATION: 310..330
; OTHER INFORMATION: Eisenberg, Schwarz, Komaromy, Wall identification method, F
; FEATURE:
; NAME/KEY: MYRISTATE
; LOCATION: 319..323
; OTHER INFORMATION: Prosite match
; FEATURE:
; NAME/KEY: MYRISTATE
; LOCATION: 323..327
; OTHER INFORMATION: Prosite match
; FEATURE:
; NAME/KEY: AMIDATION
; LOCATION: 329
; OTHER INFORMATION: Prosite match
; FEATURE:
; NAME/KEY: HELIX
; LOCATION: 333..353
; OTHER INFORMATION: Eisenberg, Schwarz, Komaromy, Wall identification method, F
; FEATURE:
; NAME/KEY: MYRISTATE
; LOCATION: 341..345
; OTHER INFORMATION: Prosite match
; FEATURE:
; NAME/KEY: PHOSPHORYLATION
; LOCATION: 350
; OTHER INFORMATION: potential protein kinase C, Prosite match
; US-09-218-207-4
```

```

Query Match          97.0%; Score 353; DB 4; Length 353;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 353; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```

QY 12 MRYLPSVVLGTAAPPYVLAMGVWLLSAPLPAPFYQALDDRLCYQSSVLEFFENYTG 71
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 1 MRYLPSVVLGTAAPPYVLAMGVWLLSAPLPAPFYQALDDRLCYQSSVLEFFENYTG 60
QY 72 VOILLYGDLPPKKNENTIIYLANHOSYVDWIVADILAIROMALGHVRVLEKGLWLPYGC 131
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 61 VOILLYGDLPPKKNENTIIYLANHOSYVDWIVADILAIROMALGHVRVLEKGLWLPYGC 120
QY 132 YFAOHGGIYVRSKAKNEKEMRNKLOSVDAGTPMYLVIFPEGSTRYNPEQTKVLSASQAF 191
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 121 YFAOHGGIYVRSKAKNEKEMRNKLOSVDAGTPMYLVIFPEGSTRYNPEQTKVLSASQAF 180
QY 192 AAQGLAVLKHLVTPRIKATHVAFDCMKNYLDAIDVTVVYEGKDDGGRRESPTTEFL 251
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 181 AAQGLAVLKHLVTPRIKATHVAFDCMKNYLDAIDVTVVYEGKDDGGRRESPTTEFL 240
QY 252 CEECPRIHIDRIDKDVPEOEHRMRWLHERFEIKDKMLIEFESPPDERRRKRPGRS 311
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 241 CEECPRIHIDRIDKDVPEOEHRMRWLHERFEIKDKMLIEFESPPDERRRKRPGRS 300
QY 312 VNSKLSIKKTLPSMLISGLTAGMLMTDAGRKLIVYNTWITIGTLGLWVTIKR 364
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 301 VNSKLSIKKTLPSMLISGLTAGMLMTDAGRKLIVYNTWITIGTLGLWVTIKR 353
```

```

RESULT 7
US-09-338-907-127
; Sequence 127, Application US/09338907
; Patent No. 6265546
; GENERAL INFORMATION:
```

```

; APPLICANT: Cohen, Daniel
; APPLICANT: Blumenfeld, Marta
; APPLICANT: Ilya, Chumakov
; APPLICANT: Bougueleret, Lydie
; TITLE OF INVENTION: PROSTATE CANCER GENE
; FILE REFERENCE: GENSET.18CPLCP
; CURRENT APPLICATION NUMBER: US/09/338,907
; EARLIER FILING DATE: 1999-06-23
; EARLIER APPLICATION NUMBER: 08/996,306
; EARLIER FILING DATE: 1997-12-22
; EARLIER APPLICATION NUMBER: 60/099,658
; EARLIER FILING DATE: 1998-09-09
; EARLIER APPLICATION NUMBER: 09/218,207
; EARLIER FILING DATE: 1998-12-22
; NUMBER OF SEQ ID NOS: 578
; SOFTWARE: Patent.pm
; SEQ ID NO 127
; LENGTH: 291
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: 98..103
; OTHER INFORMATION: Box II
; FEATURE:
; NAME/KEY: SITE
; LOCATION: 149..157
; OTHER INFORMATION: Box III
; US-09-338-907-127
```

```

Query Match          63.2%; Score 230; DB 4; Length 291;
Best Local Similarity 100.0%; Pred. No. 2.9e-223;
Matches 230; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```

QY 135 QHGGIYVRSKAKNEKEMRNKLOSVDAGTPMYLVIFPEGSTRYNPEQTKVLSASQAF 194
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 62 QHGGIYVRSKAKNEKEMRNKLOSVDAGTPMYLVIFPEGSTRYNPEQTKVLSASQAF 121
QY 195 RGLAVLKHLVTPRIKATHVAFDCMKNYLDAIDVTVVYEGKDDGGRRESPTTEFLCKE 254
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 122 RGLAVLKHLVTPRIKATHVAFDCMKNYLDAIDVTVVYEGKDDGGRRESPTTEFLCKE 181
QY 255 CEECPRIHIDRIDKDVPEOEHRMRWLHERFEIKDKMLIEFESPPDERRRKRPGRSVNS 314
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 182 CEECPRIHIDRIDKDVPEOEHRMRWLHERFEIKDKMLIEFESPPDERRRKRPGRSVNS 241
QY 315 KLSIKKTLPSMLISGLTAGMLMTDAGRKLIVYNTWITIGTLGLWVTIKR 364
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 242 KLSIKKTLPSMLISGLTAGMLMTDAGRKLIVYNTWITIGTLGLWVTIKR 291
```

```

RESULT 8
US-09-218-207-127
; Sequence 127, Application US/09218207
; Patent No. 6346381
; GENERAL INFORMATION:
; APPLICANT: Cohen, Daniel
; APPLICANT: Blumenfeld, Marta
; APPLICANT: Ilya, Chumakov
; APPLICANT: Bougueleret, Lydie
; TITLE OF INVENTION: Prostate cancer gene
; FILE REFERENCE: GENSET.018CPL
; CURRENT APPLICATION NUMBER: US/09/218,207
; EARLIER FILING DATE: 1998-12-22
; EARLIER APPLICATION NUMBER: 08/996,306
; EARLIER FILING DATE: 1997-12-22
; EARLIER APPLICATION NUMBER: 60/099,658
; EARLIER FILING DATE: 1998-09-09
; NUMBER OF SEQ ID NOS: 578
; SOFTWARE: Patent.pm
; SEQ ID NO 127
; LENGTH: 291
```

?	TYPE:	PRT	
?	ORGANISM:	Homo sapiens	
?	FEATURE:		
?	NAME/KEY:	SITE	
?	LOCATION:	98..103	
?	OTHER_INFORMATION:	Box	II
?	FEATURE:		
?	NAME/KEY:	SITE	
?	LOCATION:	149..157	
?	OTHER_INFORMATION:	Box	IIIII
US-09-218-207-127			

Query Match	63.2%	Score 230;	DB 4;	Length 291;
Best Local Similarity	100.0%;	Pred. No. 2.9e-223;		
Matches 230; Conservative	0;	Mismatches	0;	Indels 0; Gaps 0;

Qy	135	QHGGIIVYKRSKAEFKEMKRNKLSYVDAGIPMYLVIPESTRVNPEDTKVLSASQAAQ	194
Db	62	QHGGIIVYKRSKAEFKEMKRNKLSYVDAGIPMYLVIPESTRVNPEDTKVLSASQAAQ	121
Qy	195	RGIAVLKHLVLTPIKATIHVAEFDCKNKLDAIYDVTVVYEGKDDGGQRESPTMEFLCKE	254
Db	122	RGIAVLKHLVLTPIKATIHVAEFDCKNKLDAIYDVTVVYEGKDDGGQRESPTMEFLCKE	181
Qy	255	CPKIHIIHIDRLDKDOVEEBOEHMRMYLHERETIDKMLIEFYESPDERKKRRPGKSVNS	314
Db	182	CPKIHIIHIDRLDKDOVEEBOEHMRMYLHERETIDKMLIEFYESPDERKKRRPGKSVNS	241
Qy	315	KLSTIKTLPsMLILSGLTAGMLMDAGARKLYVNHWITGTLGLGCMLWNIKA	364
Db	242	KLSTIKTLPsMLILSGLTAGMLMDAGARKLYVNHWITGTLGLGCMLWNIKA	291

RESULT 9  
US-09-338-907-70

```

1      Sequence 70, Application US/09338907
2      Patent No. 6265546
3      GENERAL INFORMATION:
4      APPLICANT: Cohen, Daniel
5      APPLICANT: Blumenfeld, Marta
6      APPLICANT: Ilyia, Chumakov
7      APPLICANT: Bougueleret, Lydie
8      TITLE OF INVENTION: PROSTATE CANCER GENE
9      FILE REFERENCE: GENSER 18CP1CP
10     CURRENT APPLICATION NUMBER: US/09/338, 907
11     CURRENT FILING DATE: 1999-06-23
12     EARLIER APPLICATION NUMBER: 09/996, 306
13     EARLIER FILING DATE: 1997-12-22
14     EARLIER APPLICATION NUMBER: 60/099, 658
15     EARLIER FILING DATE: 1998-09-09
16     EARLIER APPLICATION NUMBER: 09/218, 207
17     EARLIER FILING DATE: 1998-12-22
18     NUMBER OF SEQ ID NOS: 578
19     SOFTWARE: Patent.pm
20     SEQ ID NO 70
21     LENGTH: 228
22     TYPE: PRT
23     ORGANISM: Homo sapiens
24     US-09-338-907-70

```

Query Match	62.68;	Score 228;	DB 4;	Length 228;
Best Local Similarity	100.0%;	Pred. No. 2.4e-221;		
Matches 228;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0

Qy	12	MRPSVVLIGTAPTYVLANGVRRLLSAPLPAEFOALDDRLCYVQSNVLPFEENYTG	71
Db	1	MRILLPSVVLIGTAPTYVLANGVRRLLSAPLPAEFOALDDRLCYVQSNVLPFEENYTG	60
Qy	72	VQILLGDLPRNKENIITYLANHOSVDVIVADIIIAIQNALGHVRYVLRBGLKMLPLYGC	131
Db	61	VQILLGDLPRNKENIITYLANHOSVDVIVADIIIAIQNALGHVRYVLRBGLKMLPLYGC	120

Qy	132	YFAOHGGIYVRSKAFENKEKRNKLTOSVDAGPMYLVIPPEGRNYPEDQTKVLSAQAF	101
Db	121	YFAOHGGIYVRSKAFENKEKRNKLTOSVDAGPMYLVIPPEGRNYPEDQTKVLSAQAF	180
Qy	192	AAQKGLAVLKHVLPRIKATHVADCKMKNYLDATYDVTVYIEGDDGG	239
Db	181	AAQKGLAVLKHVLPRIKATHVADCKMKNYLDATYDVTVYIEGDDGG	228

RESULT 10  
US-09-218-

```

1 Patent 70, Application US/09318207
2 : Sequence No.6346381
3 :
4 : GENERAL INFORMATION:
5 :
6 : APPLICANT: Cohen, Daniel
7 : APPLICANT: Blumenfeld, Martha
8 : APPLICANT: Ilyia, Chumakov
9 : APPLICANT: Bougueleret, Lydie
10 : TITLE OF INVENTION: Prostate cancer ge
11 :
12 : FILE REFERENCE: GENSET.018CP1
13 : CURRENT APPLICATION NUMBER: US/09/218,
14 : CURRENT FILING DATE: 1998-12-22
15 : EARLIER APPLICATION NUMBER: 08/996, 3061
16 : EARLIER FILING DATE: 1997-12-22
17 : EARLIER APPLICATION NUMBER: 60/099, 658
18 : EARLIER FILING DATE: 1998-09-09
19 :
20 : NUMBER OF SEQ ID NOS: 578
21 :
22 : SOFTWARE: Patent.pm
23 :
24 : SEQ ID NO 70
25 :
26 : LENGTH: 228
27 :
28 : TYPE: prt
29 :
30 : ORGANISM: Homo sapiens
31 :
32 : OS-09-218-207-70

```

Query Match	62.68;	Score 228;	DB 4;	Length 228;
Best Local Similarity	100.0%;	Pred. NO. 2.4e-221;		
Matches 228;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;

Qy	12	MRLLPSVILGTPPTYVLGAVWRLLSAFPAFPAFOALDRLCYQSMVLEFFENYTG	71
Db	1	MRYLTPSVVLGTAFTYVLGAVWRLLSAFPAFPAFOALDRLCYQSMVLEFFENYTG	60
Qy	72	VQILLYGDLPRKNENIIVLANHOSVDMIVADILAIRONALGHVRYVYLKEGLKMLPLYGC	131
Db	61	VQILLYGDLPRKNENIIVLANHOSVDMIVADILAIRONALGHVRYVYLKEGLKMLPLYGC	120
Qy	132	YFAQHGGIYVRSAKFENKEKRNKLOSTVDACTPMYLYVPEEGTRINPEQTKVLSAQAQF	191
Db	121	YFAQHGGIYVRSAKFENKEKRNKLOSTVDACTPMYLYVPEEGTRINPEQTKVLSAQAQF	180
Qy	192	AAQRLAALAKHVLTPRIKATIVAFDCMKNTLDALYDTVVYEGKDDG	239
Db	181	AAQRLAALAKHVLTPRIKATIVAFDCMKNTLDALYDTVVYEGKDDG	228

## RESULT 11

US-09-338-907-128  
 : Sequence 128, Application US/09338907  
 : Patent No. 6265546  
 :  
 : GENERAL INFORMATION:  
 :  
 : APPLICANT: Cohen, Daniel  
 : APPLICANT: Blumenfeld, Marla  
 : APPLICANT: Ilya, Chumakov  
 : APPLICANT: Bouqueleret, Lydie  
 :  
 : TITLE OF INVENTION: PROSTATE CANCER GENE  
 : FILE REFERENCE: GENSET.18C1CP  
 : CURRENT APPLICATION NUMBER: US/09/338.907  
 : CURRENT FILING DATE: 1999-06-23  
 : EARLIER APPLICATION NUMBER: 08/996,306  
 : EARLIER FILING DATE: 1997-12-22  
 : EARLIER APPLICATION NUMBER: 60/099,658

```

; EARLIER FILING DATE: 1998-09-09
; EARLIER APPLICATION NUMBER: 09/218, 207
; EARLIER FILING DATE: 1998-12-22
; NUMBER OF SEQ ID NOS: 578
; SOFTWARE: Patent.pm
; SEQ ID NO 128
; LENGTH: 261
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: 68..73
; OTHER INFORMATION: Box II
; NAME/KEY: SITE
; LOCATION: 119..127
; OTHER INFORMATION: Box III
US-09-338-907-128
```

```

Query Match          54.7%; Score 199; DB 4; Length 261;
Best Local Similarity 100.0%; Pred. No. 3.9e-192;
Matches 199; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```

QY 166 MYLVIEPEGTRYNPEQTKVLSASQAFPAORGGLAVLKHVLTLPRIKATHVAFDCMKNYLDAI 225
; |||||||
DB 63 MYLVIEPEGTRYNPEQTKVLSASQAFPAORGGLAVLKHVLTLPRIKATHVAFDCMKNYLDAI 122
; |||||||
QY 226 YDVTVYVEGKDDGGORRESTPTMTFELCKECPKIHIDRIDKDVPEDEOHMRWLHERF 285
; |||||||
DB 123 YDVTVYVEGKDDGGORRESTPTMTFELCKECPKIHIDRIDKDVPEDEOHMRWLHERF 182
; |||||||
QY 286 EIKDKMLIEFESPDPERRRRFPKGSVNSKLSIKTLPMSLISGLTAGMLMTDAGRKL 345
; |||||||
DB 183 EIKDKMLIEFESPDPERRRRFPKGSVNSKLSIKTLPMSLISGLTAGMLMTDAGRKL 242
; |||||||
QY 346 VNTWIVGTLLGCLMTVITKA 364
; |||||||
DB 243 VNTWIVGTLLGCLMTVITKA 261
; |||||||
```

```

RESULT 12
US-09-218-207-128
; Sequence 128, Application US/09218207
; Patent No. 6346381
; GENERAL INFORMATION:
; APPLICANT: Cohen, Daniel
; APPLICANT: Blumenfeld, Marita
; APPLICANT: Ilya, Chumakov
; APPLICANT: Bougueleret, Lydie
; TITLE OF INVENTION: Prostate cancer gene
; FILE REFERENCE: GENSET.018CPI
; CURRENT APPLICATION NUMBER: US/09/218, 207
; EARLIER FILING DATE: 1998-12-22
; EARLIER APPLICATION NUMBER: 08/996, 306
; EARLIER FILING DATE: 1997-12-22
; EARLIER APPLICATION NUMBER: 60/099, 658
; EARLIER FILING DATE: 1998-09-09
; NUMBER OF SEQ ID NOS: 578
; SOFTWARE: Patent.pm
; SEQ ID NO 128
; LENGTH: 261
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: 68..73
; OTHER INFORMATION: Box II
; NAME/KEY: SITE
; LOCATION: 119..127
; OTHER INFORMATION: Box III
US-09-218-207-128
```

```

Query Match          54.7%; Score 199; DB 4; Length 261;
Best Local Similarity 100.0%; Pred. No. 3.9e-192;
Matches 199; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```

QY 166 MYLVIEPEGTRYNPEQTKVLSASQAFPAORGGLAVLKHVLTLPRIKATHVAFDCMKNYLDAI 225
; |||||||
DB 63 MYLVIEPEGTRYNPEQTKVLSASQAFPAORGGLAVLKHVLTLPRIKATHVAFDCMKNYLDAI 122
; |||||||
QY 226 YDVTVYVEGKDDGGORRESTPTMTFELCKECPKIHIDRIDKDVPEDEOHMRWLHERF 285
; |||||||
DB 123 YDVTVYVEGKDDGGORRESTPTMTFELCKECPKIHIDRIDKDVPEDEOHMRWLHERF 182
; |||||||
QY 286 EIKDKMLIEFESPDPERRRRFPKGSVNSKLSIKTLPMSLISGLTAGMLMTDAGRKL 345
; |||||||
DB 183 EIKDKMLIEFESPDPERRRRFPKGSVNSKLSIKTLPMSLISGLTAGMLMTDAGRKL 242
; |||||||
QY 346 VNTWIVGTLLGCLMTVITKA 364
; |||||||
DB 243 VNTWIVGTLLGCLMTVITKA 261
; |||||||
```

```

RESULT 13
US-09-338-907-136
; Sequence 136, Application US/09338907
; Patent No. 6265546
; GENERAL INFORMATION:
; APPLICANT: Cohen, Daniel
; APPLICANT: Blumenfeld, Marita
; APPLICANT: Ilya, Chumakov
; APPLICANT: Bougueleret, Lydie
; TITLE OF INVENTION: PROSTATE CANCER GENE
; FILE REFERENCE: GENSET.18CPI
; CURRENT APPLICATION NUMBER: US/09/338, 907
; EARLIER FILING DATE: 1999-06-23
; EARLIER APPLICATION NUMBER: 08/996, 306
; EARLIER FILING DATE: 1997-12-22
; EARLIER APPLICATION NUMBER: 60/099, 658
; EARLIER FILING DATE: 1998-09-09
; EARLIER APPLICATION NUMBER: 09/218, 207
; EARLIER FILING DATE: 1998-12-22
; NUMBER OF SEQ ID NOS: 578
; SOFTWARE: Patent.pm
; SEQ ID NO 136
; LENGTH: 185
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: 81..83
; OTHER INFORMATION: Box I
; NAME/KEY: SITE
; LOCATION: 160..165
; OTHER INFORMATION: Box II
US-09-338-907-136
```

```

Query Match          50.8%; Score 185; DB 4; Length 185;
Best Local Similarity 100.0%; Pred. No. 3.3e-178;
Matches 185; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```

QY 12 MYLPSVVLGTAPTYVILAWGVWRLLSAFLPARFYQALDDRLCYVYQSMVLEFFENYTG 71
; |||||||
DB 1 MYLPSVVLGTAPTYVILAWGVWRLLSAFLPARFYQALDDRLCYVYQSMVLEFFENYTG 60
; |||||||
QY 72 VOILLYGDLPRKNENITILANHOSTYDWMIVADILAIQNALGHRVRLKEGLKMLPLYGC 131
; |||||||
DB 61 VOILLYGDLPRKNENITILANHOSTYDWMIVADILAIQNALGHRVRLKEGLKMLPLYGC 120
; |||||||
QY 132 YFAOHGIGIVKRSKFNKEMRKLQSYVDAGPMLVYFPGSTRNPEOTKVLASQAF 191
; |||||||
```

Db	121	YFAHGSIYVKRSAKFENEKEMRNKLQSYVDAGFPMYLVIFPEGRTRYNPQTKVLSAQAF	180
QY	192	AAQGG	196
Db	181	AAQGG	185

```

RESULT 14
US-09-218-207-136
: Sequence 136, Application US/09218207
: Patent No. 6346381
: GENERAL INFORMATION:
: APPLICANT: Cohen, Daniel
: APPLICANT: Blumenfeld, Marta
: APPLICANT: Ilya, Chumakov
: APPLICANT: Bouquelerec, Lydie
: TITLE OF INVENTION: Prostate cancer gene
: FILE REFERENCE: GENSET.018CP1
: CURRENT APPLICATION NUMBER: US/09/218,207
: CURRENT FILING DATE: 1998-12-22
: EARLIER APPLICATION NUMBER: 08/996,306
: EARLIER FILING DATE: 1997-12-22
: EARLIER APPLICATION NUMBER: 60/099,658
: EARLIER FILING DATE: 1998-09-09
: NUMBER OF SEQ ID NOS: 578
: SOFTWARE: Patent.pm
: SEQ ID NO 136
: LENGTH: 185
: TYPE: PRT
: ORGANISM: Homo sapiens
: FEATURE:
: NAME/KEY: SITE
: LOCATION: 81..83
: OTHER INFORMATION: Box I
: FEATURE:
: NAME/KEY: SITE
: LOCATION: 160..165
: OTHER INFORMATION: Box II
US-09-218-207-136

```

```

CURRENT APPLICATION NUMBER: US-09/338,907
CURRENT FILING DATE: 1999-06-23
EARLIER APPLICATION NUMBER: 08/996,306
EARLIER FILING DATE: 1997-12-22
EARLIER APPLICATION NUMBER: 60/099,658
EARLIER FILING DATE: 1998-09-09
EARLIER APPLICATION NUMBER: 09/218,207
EARLIER FILING DATE: 1998-12-22
NUMBER OF SEQ ID NOS: 578
SOFTWARE: Patent.pm
SEQ ID NO 134
LENGTH: 315
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: SITE
LOCATION: 81..83
OTHER INFORMATION: Box I
FEATURE:
NAME/KEY: SITE
LOCATION: 160..165
OTHER INFORMATION: Box II
US-09-338-907-134

```

```
Search completed: August 28, 2002, 11:19:21
Job time: 304 sec
```

RESULT 15  
 US-09-338-907-134  
 : Sequence 134, Application US/09338907  
 : Patient No. 6265546  
 : GENERAL INFORMATION:  
 : APPLICANT: Cohen, Daniel  
 : APPLICANT: Blumenfeld, Maita  
 : APPLICANT: Ilya, Chumakov  
 : APPLICANT: Bouguetelert, Lydie  
 : TITLE OF INVENTION: PROSTATE CANCER GENE  
 : FILE REFERENCE: GENSET.18C1CIP

11.

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 28, 2002, 11:25:19 ; Search time 305.3 Seconds  
(without alignments)  
419.656 Million cell updates/sec

Title: US-09-853-526-5

Perfect score: 364  
Sequence: 1 MLSTLVHTYSMRYLPSVY.....YVNTWLYGFLGLGLMTYTIKA 364

Scoring table:

Gapop 60.0 , Gapext 60.0

Searched: 3502263 seqs, 351980561 residues

Word size: 0

Total number of hits satisfying chosen parameters: 3502263

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database: Pending\_Patents\_AA\_Main:\*

- 1: /cgn2\_6/ptodata/2/paa/PCYUS\_COMB.pep:\*
- 2: /cgn2\_6/ptodata/2/paa/US06\_COMB.pep:\*
- 3: /cgn2\_6/ptodata/2/paa/US07\_COMB.pep:\*
- 4: /cgn2\_6/ptodata/2/paa/US080\_COMB.pep:\*
- 5: /cgn2\_6/ptodata/2/paa/US081\_COMB.pep:\*
- 6: /cgn2\_6/ptodata/2/paa/US082\_COMB.pep:\*
- 7: /cgn2\_6/ptodata/2/paa/US083\_COMB.pep:\*
- 8: /cgn2\_6/ptodata/2/paa/US084\_COMB.pep:\*
- 9: /cgn2\_6/ptodata/2/paa/US085\_COMB.pep:\*
- 10: /cgn2\_6/ptodata/2/paa/US086\_COMB.pep:\*
- 11: /cgn2\_6/ptodata/2/paa/US087\_COMB.pep:\*
- 12: /cgn2\_6/ptodata/2/paa/US088\_COMB.pep:\*
- 13: /cgn2\_6/ptodata/2/paa/US089\_COMB.pep:\*
- 14: /cgn2\_6/ptodata/2/paa/US090\_COMB.pep:\*
- 15: /cgn2\_6/ptodata/2/paa/US091\_COMB.pep:\*
- 16: /cgn2\_6/ptodata/2/paa/US092\_COMB.pep:\*
- 17: /cgn2\_6/ptodata/2/paa/US093\_COMB.pep:\*
- 18: /cgn2\_6/ptodata/2/paa/US094\_COMB.pep:\*
- 19: /cgn2\_6/ptodata/2/paa/US095\_COMB.pep:\*
- 20: /cgn2\_6/ptodata/2/paa/US096\_COMB.pep:\*
- 21: /cgn2\_6/ptodata/2/paa/US097\_COMB.pep:\*
- 22: /cgn2\_6/ptodata/2/paa/US098\_COMB.pep:\*
- 23: /cgn2\_6/ptodata/2/paa/US099\_COMB.pep:\*
- 24: /cgn2\_6/ptodata/2/paa/US100\_COMB.pep:\*
- 25: /cgn2\_6/ptodata/2/paa/US101\_COMB.pep:\*
- 26: /cgn2\_6/ptodata/2/paa/US60\_COMB.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	364	100.0	364	US-09-436-919-1	Sequence 1, Appli
2	364	100.0	364	US-09-817-910-7	Sequence 7, Appli
3	364	100.0	364	US-09-853-526-5	Sequence 5, Appli
4	364	100.0	364	US-09-901-484A-5	Sequence 5, Appli
5	364	100.0	364	US-60-099-658-5	Sequence 5, Appli
6	353	97.0	353	PCT-US01-01431-59	Sequence 59, Appli
7	353	97.0	353	PCT-US01-01431-79	Sequence 79, Appli

8	353	97.0	353	1	PCT-US01-11988-1474	Sequence 1474, Ap
9	353	97.0	353	1	PCT-US01-11988-1475	Sequence 1475, Ap
10	353	97.0	353	22	US-09-833-245-1474	Sequence 1474, Ap
11	353	97.0	353	22	US-09-833-245-1475	Sequence 1475, Ap
12	353	97.0	353	22	US-09-853-526-4	Sequence 4, Appli
13	353	97.0	353	23	US-09-901-484A-4	Sequence 4, Appli
14	353	97.0	353	23	US-09-915-582-59	Sequence 59, Appli
15	353	97.0	353	23	US-09-915-582-79	Sequence 79, Appli
16	353	97.0	353	26	US-60-099-658-4	Sequence 4, Appli
17	290	79.7	353	18	US-09-488-275A-2736	Sequence 2736, Ap
18	269	73.9	269	1	PCT-US01-01327-47	Sequence 47, Appli
19	263	72.3	372	18	US-09-488-725A-6308	Sequence 6308, Ap
20	248	68.1	450	26	US-60-245-222-110	Sequence 110, App
21	230	63.2	291	22	US-09-853-526-127	Sequence 127, App
22	228	62.6	228	22	US-09-901-484A-127	Sequence 127, App
23	228	62.6	228	22	US-09-853-526-70	Sequence 70, Appli
24	228	62.6	228	23	US-09-901-484A-70	Sequence 70, Appli
25	228	62.6	228	26	US-60-099-658-70	Sequence 128, App
26	199	54.7	261	22	US-09-853-526-128	Sequence 128, App
27	199	54.7	261	23	US-09-901-484A-128	Sequence 128, App
28	185	50.8	185	22	US-09-853-526-136	Sequence 136, App
29	185	50.8	185	23	US-09-901-484A-136	Sequence 136, App
30	185	50.8	185	22	US-09-853-526-134	Sequence 134, App
31	185	50.8	315	23	US-09-901-484A-134	Sequence 134, App
32	184	50.5	300	22	US-09-853-526-135	Sequence 135, App
33	184	50.5	300	23	US-09-901-484A-135	Sequence 135, App
34	154	42.3	182	22	US-09-853-526-133	Sequence 133, App
35	154	42.3	182	23	US-09-901-484A-133	Sequence 133, App
36	153	42.0	176	18	US-09-436-919-5	Sequence 5, Appli
37	151	41.5	1032	26	US-60-212-413-309	Sequence 309, App
38	151	41.5	1032	26	US-60-229-518-248	Sequence 248, App
39	116	31.9	238	22	US-09-853-526-126	Sequence 126, App
40	116	31.9	238	23	US-09-901-484A-126	Sequence 126, App
41	85	23.4	97	22	US-09-853-526-132	Sequence 132, App
42	85	23.4	97	23	US-09-901-484A-132	Sequence 132, App
43	81	22.3	92	17	US-09-316-123-124	Sequence 124, App
44	81	22.3	92	17	US-09-827-244-124	Sequence 124, App
45	70	19.2	257	1	PCT-US01-08631-31719	Sequence 31719, A

ALIGNMENTS

RESULT 1  
US-09-436-919-1  
; Sequence 1, Application US/09436919A  
; GENERAL INFORMATION:  
; APPLICANT: Leung, David W  
; TITLE OF INVENTION: Human Lysophosphatidic Acid Acyltransferase-epsilon  
; FILE REFERENCE: 1801B  
; CURRENT APPLICATION NUMBER: US/09/436, 919A  
; CURRENT FILING DATE: 1999-11-09  
; NUMBER OF SEQ ID NOS: 5  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 1  
; LENGTH: 364  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; FEATURE:  
; OTHER INFORMATION: hLPAATepsilon  
US-09-436-919-1

Query Match	100.0%	Score 364:	DB 18:	Length 364:
Best Local Similarity	100.0%	Pred. No. 0:		
Matches 364:	Conservative	0:	Mismatches	0:
Indels	0:	Gaps	0:	
QY	1	MLSTLVHTYSMRYLPSVVLGTAPTYLVANGVWLLSAPLPARYOALDRLCYVOS	60	
DB	1	MLSTLVHTYSMRYLPSVVLGTAPTYLVANGVWLLSAPLPARYOALDRLCYVOS	60	
QY	61	MYLFEENTGVQILYGLDLPKNKENIIVLANHSTVDIVADIAIRONALGHVRYVK	120	

Db 61 MVLFFENNTGVOILLYGDLPKNKENIITYLANHOSTVDMIVADILAIROMALGHVRYVLK 120  
QY 121 EGLKMLPLYGCVFAOHGIIYVKSASFNEKEMNKLOSVDAGTPLYVIFPEGTRYNPE 180  
Db 121 EGLKMLPLYGCVFAOHGIIYVKSASFNEKEMNKLOSVDAGTPLYVIFPEGTRYNPE 180  
QY 181 QTKVLSASQAFAAORGLAVLKHVLPRIKATHVAFDCMKNYLDATDVVYVEGKDDGQ 240  
Db 181 QTKVLSASQAFAAORGLAVLKHVLPRIKATHVAFDCMKNYLDATDVVYVEGKDDGQ 240  
QY 241 RRESPTMTFELCKECPKIHIDRIKDDVPEOEHRMRHLHERFEIKDKMLIEFYESPD 300  
Db 241 RRESPTMTFELCKECPKIHIDRIKDDVPEOEHRMRHLHERFEIKDKMLIEFYESPD 300  
QY 301 PERRRRPCKSVNSKLSIKKTLPSMLISGLTAGMLMTDAGRKLVTWYTWIYGTLLGCLMW 360  
Db 301 PERRRRPCKSVNSKLSIKKTLPSMLISGLTAGMLMTDAGRKLVTWYTWIYGTLLGCLMW 360  
QY 361 TTKA 364  
Db 361 TTKA 364

RESULT 2  
US-09-817-910-7  
; Sequence 7, Application US/09817910  
; GENERAL INFORMATION:  
; APPLICANT: Meyers, Rachel A.  
; APPLICANT: Macbeth, Kyle J.  
; APPLICANT: Williamson, Mark  
; APPLICANT: Rudolph-Owen, Laura A.  
; APPLICANT: Tsai, Fong-Ying  
; TITLE OF INVENTION: 46743 and 27417 NOVEL HUMAN  
; FILE REFERENCE: 10448-032001  
; CURRENT APPLICATION NUMBER: US/09/817,910  
; CURRENT FILING DATE: 2001-03-26  
; PRIOR APPLICATION NUMBER: US 60/192,092  
; PRIOR FILING DATE: 2000-03-24  
; NUMBER OF SEQ ID NOS: 15  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 7  
; LENGTH: 364  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-817-910-7

Query Match 100.0%; Score 364; DB 22; Length 364;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 364; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 MVLSTLVHTYSMRYLLPSVVLGTAPTYVLAWGVWLLSAFLPARYOALDDRLXYCVOS 60  
Db 1 MVLSTLVHTYSMRYLLPSVVLGTAPTYVLAWGVWLLSAFLPARYOALDDRLXYCVOS 60  
QY 61 MVLFFENTGVQIILYGDLPKNKENIITYLANHOSTVDMIVADILAIROMALGHVRYVLK 120  
Db 61 MVLFFENTGVQIILYGDLPKNKENIITYLANHOSTVDMIVADILAIROMALGHVRYVLK 120  
QY 121 EGLKMLPLYGCVFAOHGIIYVKSASFNEKEMNKLOSVDAGTPLYVIFPEGTRYNPE 180  
Db 121 EGLKMLPLYGCVFAOHGIIYVKSASFNEKEMNKLOSVDAGTPLYVIFPEGTRYNPE 180  
QY 181 QTKVLSASQAFAAORGLAVLKHVLPRIKATHVAFDCMKNYLDATDVVYVEGKDDGQ 240  
Db 181 QTKVLSASQAFAAORGLAVLKHVLPRIKATHVAFDCMKNYLDATDVVYVEGKDDGQ 240  
QY 241 RRESPTMTFELCKECPKIHIDRIKDDVPEOEHRMRHLHERFEIKDKMLIEFYESPD 300  
Db 241 RRESPTMTFELCKECPKIHIDRIKDDVPEOEHRMRHLHERFEIKDKMLIEFYESPD 300  
QY 301 PERRRRPCKSVNSKLSIKKTLPSMLISGLTAGMLMTDAGRKLVTWYTWIYGTLLGCLMW 360  
Db 301 PERRRRPCKSVNSKLSIKKTLPSMLISGLTAGMLMTDAGRKLVTWYTWIYGTLLGCLMW 360

Db 301 PERRRRPCKSVNSKLSIKKTLPSMLISGLTAGMLMTDAGRKLVTWYTWIYGTLLGCLMW 360  
QY 361 TTKA 364  
Db 361 TTKA 364

RESULT 3  
US-09-853-526-5  
; Sequence 5, Application US/09853526  
; GENERAL INFORMATION:  
; APPLICANT: Cohen, Daniel  
; APPLICANT: Blumenfeld, Marta  
; APPLICANT: Ilya, Chumakov  
; APPLICANT: Bougueleret, Lydie  
; TITLE OF INVENTION: PROSTATE CANCER GENE  
; FILE REFERENCE: GENSET.18C1PC  
; CURRENT APPLICATION NUMBER: US/09/853,526  
; CURRENT FILING DATE: 2001-05-11  
; PRIOR APPLICATION NUMBER: 09/338,907  
; PRIOR FILING DATE: 1999-06-23  
; PRIOR APPLICATION NUMBER: 08/996,306  
; PRIOR FILING DATE: 1997-12-22  
; PRIOR APPLICATION NUMBER: 60/099,658  
; PRIOR FILING DATE: 1998-09-09  
; PRIOR APPLICATION NUMBER: 09/218,207  
; PRIOR FILING DATE: 1998-12-22  
; NUMBER OF SEQ ID NOS: 578  
; SOFTWARE: Patent.pm  
; SEQ ID NO 5  
; LENGTH: 364  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-853-526-5

Query Match 100.0%; Score 364; DB 22; Length 364;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 364; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 MVLSTLVHTYSMRYLLPSVVLGTAPTYVLAWGVWLLSAFLPARYOALDDRLXYCVOS 60  
Db 1 MVLSTLVHTYSMRYLLPSVVLGTAPTYVLAWGVWLLSAFLPARYOALDDRLXYCVOS 60  
QY 61 MVLFFENTGVQIILYGDLPKNKENIITYLANHOSTVDMIVADILAIROMALGHVRYVLK 120  
Db 61 MVLFFENTGVQIILYGDLPKNKENIITYLANHOSTVDMIVADILAIROMALGHVRYVLK 120  
QY 121 EGLKMLPLYGCVFAOHGIIYVKSASFNEKEMNKLOSVDAGTPLYVIFPEGTRYNPE 180  
Db 121 EGLKMLPLYGCVFAOHGIIYVKSASFNEKEMNKLOSVDAGTPLYVIFPEGTRYNPE 180  
QY 181 QTKVLSASQAFAAORGLAVLKHVLPRIKATHVAFDCMKNYLDATDVVYVEGKDDGQ 240  
Db 181 QTKVLSASQAFAAORGLAVLKHVLPRIKATHVAFDCMKNYLDATDVVYVEGKDDGQ 240  
QY 241 RRESPTMTFELCKECPKIHIDRIKDDVPEOEHRMRHLHERFEIKDKMLIEFYESPD 300  
Db 241 RRESPTMTFELCKECPKIHIDRIKDDVPEOEHRMRHLHERFEIKDKMLIEFYESPD 300  
QY 301 PERRRRPCKSVNSKLSIKKTLPSMLISGLTAGMLMTDAGRKLVTWYTWIYGTLLGCLMW 360  
Db 301 PERRRRPCKSVNSKLSIKKTLPSMLISGLTAGMLMTDAGRKLVTWYTWIYGTLLGCLMW 360  
QY 361 TTKA 364  
Db 361 TTKA 364

RESULT 4  
US-09-901-484A-5  
; Sequence 5, Application US/09901484A



```

; GENERAL INFORMATION:
; APPLICANT: Cohen, Daniel
; APPLICANT: Blumenfeld, Marta
; APPLICANT: Chumakov, Ilya
; APPLICANT: Bougueleret, Lydie
; TITLE OF INVENTION: Prostate Cancer Gene
; FILE REFERENCE: GEN-111XC3D2
; CURRENT APPLICATION NUMBER: US/09/901,484A
; CURRENT FILING DATE: 2001-07-09
; PRIOR APPLICATION NUMBER: US 08/996,306
; PRIOR FILING DATE: 1997-12-22
; PRIOR APPLICATION NUMBER: US 60/099,658
; PRIOR FILING DATE: 1998-09-09
; PRIOR APPLICATION NUMBER: US 09/218,207
; PRIOR FILING DATE: 1998-12-22
; PRIOR APPLICATION NUMBER: US 09/338,907
; PRIOR FILING DATE: 1999-06-23
; PRIOR APPLICATION NUMBER: US 09/853,526
; PRIOR FILING DATE: 2001-05-11
; NUMBER OF SEQ ID NOS: 578
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 5
; LENGTH: 364
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-901-484A-5

```

```

Query Match      100.0%; Score 364; DB 23; Length 364;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 364; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1 MLTSLVHTYSMRYLLPSVVLGTAPTYYVLAMGVWRLLSAFLPARFYQALDDRLCYGQS 60
DB 1 MLTSLVHTYSMRYLLPSVVLGTAPTYYVLAMGVWRLLSAFLPARFYQALDDRLCYGQS 60
QY 61 MYLFFENTYGVQILLYGDLPRKNENIYLANHOSYVDWIVADILAIRONALGHRVYVK 120
DB 61 MYLFFENTYGVQILLYGDLPRKNENIYLANHOSYVDWIVADILAIRONALGHRVYVK 120
QY 121 EGIKWLPLYGCTYFAOHGCIYVKSAPFNEKEMRNKLOSVDAGTPLYVTFPEGSTRYNPE 180
DB 121 EGIKWLPLYGCTYFAOHGCIYVKSAPFNEKEMRNKLOSVDAGTPLYVTFPEGSTRYNPE 180
QY 181 QTVLASASQAFQAORGLAVLKHLTPRIKATHVAFDCMKNYLDAIYDVTVVYEGKDDGGQ 240
DB 181 QTVLASASQAFQAORGLAVLKHLTPRIKATHVAFDCMKNYLDAIYDVTVVYEGKDDGGQ 240
QY 241 RRESPTMTFELCKECPKIHIDRIDKDVPEOEHRRLHREFEIKDKMLIEFYESPD 300
DB 241 RRESPTMTFELCKECPKIHIDRIDKDVPEOEHRRLHREFEIKDKMLIEFYESPD 300
QY 301 PERRRRPFGKSVNSKLSIKKTLPSMLILSGLTAGMLMTDAGRKLTVNTWIYGTLLGCLWV 360
DB 301 PERRRRPFGKSVNSKLSIKKTLPSMLILSGLTAGMLMTDAGRKLTVNTWIYGTLLGCLWV 360
QY 361 TIRA 364
DB 361 TIRA 364

```

```

RESULT 5
US-60-099-658-5
; Sequence 5, Application US/60099658
; GENERAL INFORMATION:
; APPLICANT: Cohen, Daniel
; APPLICANT: Chumakov, Ilya
; APPLICANT: Blumenfeld, Marta
; APPLICANT: Bougueleret, Lydie
; TITLE OF INVENTION: Prostate cancer gene
; NUMBER OF SEQUENCES: 99
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Knobbé, Martens, Olson & Bear

```

```

; STREET: 501 West Broadway
; City: San Diego
; STATE: California
; COUNTRY: USA
; ZIP: 92101-3505
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: win95
; SOFTWARE: word
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/60/099,658
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Israelson, Ned A.
; REGISTRATION NUMBER: 29,655
; REFERENCE/DOCKET NUMBER: GENSET.018APR
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (619) 235-8550
; TELEFAX: (619) 235-0176
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 364 amino acids
; TYPE: AMINO ACID
; STRANDEDNESS: SINGLE
; TOPOLOGY: LINEAR
; MOLECULE TYPE: PROTEIN
; ORIGINAL SOURCE:
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: potential protein
; LOCATION: 1..364
US-60-099-658-5

```

```

Query Match      100.0%; Score 364; DB 26; Length 364;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 364; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1 MLTSLVHTYSMRYLLPSVVLGTAPTYYVLAMGVWRLLSAFLPARFYQALDDRLCYGQS 60
DB 1 MLTSLVHTYSMRYLLPSVVLGTAPTYYVLAMGVWRLLSAFLPARFYQALDDRLCYGQS 60
QY 61 MYLFFENTYGVQILLYGDLPRKNENIYLANHOSYVDWIVADILAIRONALGHRVYVK 120
DB 61 MYLFFENTYGVQILLYGDLPRKNENIYLANHOSYVDWIVADILAIRONALGHRVYVK 120
QY 121 EGIKWLPLYGCTYFAOHGCIYVKSAPFNEKEMRNKLOSVDAGTPLYVTFPEGSTRYNPE 180
DB 121 EGIKWLPLYGCTYFAOHGCIYVKSAPFNEKEMRNKLOSVDAGTPLYVTFPEGSTRYNPE 180
QY 181 QTVLASASQAFQAORGLAVLKHLTPRIKATHVAFDCMKNYLDAIYDVTVVYEGKDDGGQ 240
DB 181 QTVLASASQAFQAORGLAVLKHLTPRIKATHVAFDCMKNYLDAIYDVTVVYEGKDDGGQ 240
QY 241 RRESPTMTFELCKECPKIHIDRIDKDVPEOEHRRLHREFEIKDKMLIEFYESPD 300
DB 241 RRESPTMTFELCKECPKIHIDRIDKDVPEOEHRRLHREFEIKDKMLIEFYESPD 300
QY 301 PERRRRPFGKSVNSKLSIKKTLPSMLILSGLTAGMLMTDAGRKLTVNTWIYGTLLGCLWV 360
DB 301 PERRRRPFGKSVNSKLSIKKTLPSMLILSGLTAGMLMTDAGRKLTVNTWIYGTLLGCLWV 360
QY 361 TIRA 364
DB 361 TIRA 364

```

```

RESULT 6
PCT-US01-01431-59
; Sequence 59, Application PC/TUS0101431
; GENERAL INFORMATION:

```



```

QY 252 CKECPKIHIDRIDKKDVEEEDHMRRLHEFEIKDKMLIEFYSPDERRRKPPGKS 311
      |||||||
Db 241 CKECPKIHIDRIDKKDVEEEDHMRRLHEFEIKDKMLIEFYSPDERRRKPPGKS 300
      |||||||
QY 312 VNSKLSIKKTLPSSMLISGLTAGMLMTDAGRKLYVTWTWYIGTLGCLWYTIKA 364
      |||||||
Db 301 VNSKLSIKKTLPSSMLISGLTAGMLMTDAGRKLYVTWTWYIGTLGCLWYTIKA 353

```

```

RESULT 9
PCT-US01-11988-1475
; Sequence 1475, Application PC/TUS0111988
; GENERAL INFORMATION:
; APPLICANT: Human Genome Sciences, Inc.
; TITLE OF INVENTION: Albumin Fusion Proteins
; FILE REFERENCE: PF546PCT
; CURRENT APPLICATION NUMBER: PCT/US01/11988
; CURRENT FILING DATE: 2001-01-12
; PRIOR APPLICATION NUMBER: 60/229, 358
; PRIOR FILING DATE: 2000-04-12
; PRIOR APPLICATION NUMBER: 60/256, 931
; PRIOR FILING DATE: 2000-12-21
; PRIOR APPLICATION NUMBER: 60/199, 384
; PRIOR FILING DATE: 2000-04-25
; NUMBER OF SEQ ID NOS: 2267
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1475
; LENGTH: 353
; TYPE: PRT
; ORGANISM: Homo sapiens
PCT-US01-11988-1475

```

```

Query Match          97.0%: Score 353; DB 1; Length 353;
Best Local Similarity 100.0%: Pred. No. 0;
Matches 353; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 12 MRYLPSVVLGTAPTYVLAMGWRLLSAFLPARFYQALDDRLCYQSWLFEFFENYTG 71
      |||||||
Db 1 MRYLPSVVLGTAPTYVLAMGWRLLSAFLPARFYQALDDRLCYQSWLFEFFENYTG 60
      |||||||
QY 72 VOILLYGDLPKKNENITITLANHSTVDWIVADILAIQNALGHVRYLKEGLKWLPLYGC 131
      |||||||
Db 61 VOILLYGDLPKKNENITITLANHSTVDWIVADILAIQNALGHVRYLKEGLKWLPLYGC 120
      |||||||
QY 132 YFAOHGIIYVKRSKAFNEKEMRKLOSVDAGTPMTLVIFPESTRYNPEOTKVLASQAF 191
      |||||||
Db 121 YFAOHGIIYVKRSKAFNEKEMRKLOSVDAGTPMTLVIFPESTRYNPEOTKVLASQAF 180
      |||||||
QY 192 AAORGLAVLKHVLTPIKATHVAFDCMKNYLDAIYDVTVYVEEGKDDGGORRESPTWTEFL 251
      |||||||
Db 181 AAORGLAVLKHVLTPIKATHVAFDCMKNYLDAIYDVTVYVEEGKDDGGORRESPTWTEFL 240
      |||||||
QY 252 CKECPKIHIDRIDKKDVEEEDHMRRLHEFEIKDKMLIEFYSPDERRRKPPGKS 311
      |||||||
Db 241 CKECPKIHIDRIDKKDVEEEDHMRRLHEFEIKDKMLIEFYSPDERRRKPPGKS 300
      |||||||
QY 312 VNSKLSIKKTLPSSMLISGLTAGMLMTDAGRKLYVTWTWYIGTLGCLWYTIKA 364
      |||||||
Db 301 VNSKLSIKKTLPSSMLISGLTAGMLMTDAGRKLYVTWTWYIGTLGCLWYTIKA 353

```

```

RESULT 10
US-09-833-245-1474
; Sequence 1474, Application US/09833245
; GENERAL INFORMATION:
; APPLICANT: Human Genome Sciences, Inc.
; TITLE OF INVENTION: Albumin Fusion Proteins
; FILE REFERENCE: PF546PCT
; CURRENT APPLICATION NUMBER: US/09/833,245
; CURRENT FILING DATE: 2001-04-12
; PRIOR APPLICATION NUMBER: 60/229, 358
; PRIOR FILING DATE: 2000-04-12

```

```

; PRIOR APPLICATION NUMBER: 60/256, 931
; PRIOR FILING DATE: 2000-12-21
; PRIOR APPLICATION NUMBER: 60/199, 384
; PRIOR FILING DATE: 2000-04-25
; NUMBER OF SEQ ID NOS: 2267
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1474
; LENGTH: 353
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-833-245-1474

```

```

Query Match          97.0%: Score 353; DB 22; Length 353;
Best Local Similarity 100.0%: Pred. No. 0;
Matches 353; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 12 MRYLPSVVLGTAPTYVLAMGWRLLSAFLPARFYQALDDRLCYQSWLFEFFENYTG 71
      |||||||
Db 1 MRYLPSVVLGTAPTYVLAMGWRLLSAFLPARFYQALDDRLCYQSWLFEFFENYTG 60
      |||||||
QY 72 VOILLYGDLPKKNENITITLANHSTVDWIVADILAIQNALGHVRYLKEGLKWLPLYGC 131
      |||||||
Db 61 VOILLYGDLPKKNENITITLANHSTVDWIVADILAIQNALGHVRYLKEGLKWLPLYGC 120
      |||||||
QY 132 YFAOHGIIYVKRSKAFNEKEMRKLOSVDAGTPMTLVIFPESTRYNPEOTKVLASQAF 191
      |||||||
Db 121 YFAOHGIIYVKRSKAFNEKEMRKLOSVDAGTPMTLVIFPESTRYNPEOTKVLASQAF 180
      |||||||
QY 192 AAORGLAVLKHVLTPIKATHVAFDCMKNYLDAIYDVTVYVEEGKDDGGORRESPTWTEFL 251
      |||||||
Db 181 AAORGLAVLKHVLTPIKATHVAFDCMKNYLDAIYDVTVYVEEGKDDGGORRESPTWTEFL 240
      |||||||
QY 252 CKECPKIHIDRIDKKDVEEEDHMRRLHEFEIKDKMLIEFYSPDERRRKPPGKS 311
      |||||||
Db 241 CKECPKIHIDRIDKKDVEEEDHMRRLHEFEIKDKMLIEFYSPDERRRKPPGKS 300
      |||||||
QY 312 VNSKLSIKKTLPSSMLISGLTAGMLMTDAGRKLYVTWTWYIGTLGCLWYTIKA 364
      |||||||
Db 301 VNSKLSIKKTLPSSMLISGLTAGMLMTDAGRKLYVTWTWYIGTLGCLWYTIKA 353

```

```

RESULT 11
US-09-833-245-1475
; Sequence 1475, Application US/09833245
; GENERAL INFORMATION:
; APPLICANT: Human Genome Sciences, Inc.
; TITLE OF INVENTION: Albumin Fusion Proteins
; FILE REFERENCE: PF546PCT
; CURRENT APPLICATION NUMBER: US/09/833,245
; CURRENT FILING DATE: 2001-04-12
; PRIOR APPLICATION NUMBER: 60/229, 358
; PRIOR FILING DATE: 2000-04-12
; PRIOR APPLICATION NUMBER: 60/256, 931
; PRIOR FILING DATE: 2000-12-21
; PRIOR APPLICATION NUMBER: 60/199, 384
; PRIOR FILING DATE: 2000-04-25
; NUMBER OF SEQ ID NOS: 2267
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1475
; LENGTH: 353
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-833-245-1475

```

```

Query Match          97.0%: Score 353; DB 22; Length 353;
Best Local Similarity 100.0%: Pred. No. 0;
Matches 353; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 72 VOILLYGDLPKNNENIYLANHSTVDWIVADILAIROMALGHVRYVLEKGLKMLPLYGC 131  
DB 61 VOILLYGDLPKNNENIYLANHSTVDWIVADILAIROMALGHVRYVLEKGLKMLPLYGC 120  
QY 132 YFADHGGIYVRSKAKFEKEMKRNKLOSIVDAGTPMYIYIEPEGRIYRPEOTKVLASASQAF 191  
DB 121 YFADHGGIYVRSKAKFEKEMKRNKLOSIVDAGTPMYIYIEPEGRIYRPEOTKVLASASQAF 180  
QY 192 AAOGLAVLKHVLTLPRIKATHVAFDCMKNYLDAIYDVTVYEGKDDGSGORRESPTMEFL 251  
DB 181 AAOGLAVLKHVLTLPRIKATHVAFDCMKNYLDAIYDVTVYEGKDDGSGORRESPTMEFL 240  
QY 252 CKECPKIHIDRIKDKVPEEOEHMRMLHEREFIDKMLIEFYESPDEPRKRFPKGS 311  
DB 241 CKECPKIHIDRIKDKVPEEOEHMRMLHEREFIDKMLIEFYESPDEPRKRFPKGS 300  
QY 312 VNSKLSIKTKLPMSLLISGLTAGMLMDAGRKLYVNTWYIGTLGCLMWTIKA 364  
DB 301 VNSKLSIKTKLPMSLLISGLTAGMLMDAGRKLYVNTWYIGTLGCLMWTIKA 353

## RESULT 12

US-09-853-526-4

Sequence 4, Application US/09853526

GENERAL INFORMATION:

APPLICANT: Cohen, Daniel

APPLICANT: Blumenfeld, Marta

APPLICANT: Ilyu, Chumakov

APPLICANT: Bouguenelat, Lydie

TITLE OF INVENTION: PROSTATE CANCER GENE

FILE REFERENCE: GENSET.18CP1CP

CURRENT APPLICATION NUMBER: US/09/853,526

PRIOR FILING DATE: 2001-05-11

PRIOR APPLICATION NUMBER: 09/338,907

PRIOR FILING DATE: 1999-06-23

PRIOR APPLICATION NUMBER: 08/996,306

PRIOR FILING DATE: 1997-12-22

PRIOR APPLICATION NUMBER: 60/099,658

PRIOR FILING DATE: 1998-09-09

PRIOR APPLICATION NUMBER: 09/218,207

PRIOR FILING DATE: 1998-12-22

NUMBER OF SEQ ID NOS: 578

SOFTWARE: Patent.pm

SEQ ID NO 4

LENGTH: 353

TYPE: PRT

ORGANISM: Homo sapiens

FEATURE:

NAME/KEY: HELIX

LOCATION: 1..33

OTHER INFORMATION: Rao and Argos identification method, potential helix

NAME/KEY: HELIX

LOCATION: 4..20

OTHER INFORMATION: Klein, Kanehisa and Delisi identification method, potential

NAME/KEY: HELIX

LOCATION: 4..24

OTHER INFORMATION: Eisenberg, Schwarz, Komarony, Wall identification method, F

NAME/KEY: MYRISTATE

LOCATION: 12..16

OTHER INFORMATION: Prosite match

NAME/KEY: HELIX

LOCATION: 50..70

OTHER INFORMATION: Eisenberg, Schwarz, Komarony, Wall identification method, F

NAME/KEY: CARBOHYD

LOCATION: 37..59

OTHER INFORMATION: Prosite match

NAME/KEY: HELIX

LOCATION: 76..96

OTHER INFORMATION: Eisenberg, Schwarz, Komarony, Wall identification method, F

OTHER INFORMATION: helix

NAME/KEY: PHOSPHORYLATION  
LOCATION: 78  
OTHER INFORMATION: potential Tyrosine kinase site, Prosite match  
NAME/KEY: PHOSPHORYLATION  
LOCATION: 84  
OTHER INFORMATION: potential caseine kinase II site, Prosite match  
NAME/KEY: SITE  
LOCATION: 94..115  
OTHER INFORMATION: potential Leucine zipper site, Prosite match  
NAME/KEY: MYRISTATE  
LOCATION: 119..123  
OTHER INFORMATION: potential site, Prosite match  
NAME/KEY: PHOSPHORYLATION  
LOCATION: 133  
OTHER INFORMATION: potential protein kinase C, Prosite match  
NAME/KEY: PHOSPHORYLATION  
LOCATION: 147  
OTHER INFORMATION: potential caseine kinase II site, Prosite match  
NAME/KEY: PHOSPHORYLATION  
LOCATION: 194  
OTHER INFORMATION: potential protein kinase C, Prosite match  
NAME/KEY: PHOSPHORYLATION  
LOCATION: 215  
OTHER INFORMATION: potential Tyrosine kinase site, Prosite match  
NAME/KEY: SULFATATION  
LOCATION: 221  
OTHER INFORMATION: Prosite match  
NAME/KEY: PHOSPHORYLATION  
LOCATION: 233  
OTHER INFORMATION: potential cAMP and cGMP dependant protein kinase site,  
OTHER INFORMATION: match  
NAME/KEY: PHOSPHORYLATION  
LOCATION: 235  
OTHER INFORMATION: potential caseine kinase II site, Prosite match  
NAME/KEY: PHOSPHORYLATION  
LOCATION: 306  
OTHER INFORMATION: potential protein kinase C, Prosite match  
NAME/KEY: HELIX  
LOCATION: 310..330  
OTHER INFORMATION: Eisenberg, Schwarz, Komarony, Wall identification method,  
OTHER INFORMATION: helix  
NAME/KEY: MYRISTATE  
LOCATION: 319..323  
OTHER INFORMATION: Prosite match  
NAME/KEY: MYRISTATE  
LOCATION: 323..327  
OTHER INFORMATION: Prosite match  
NAME/KEY: AMIDATION  
LOCATION: 329  
OTHER INFORMATION: Prosite match  
NAME/KEY: HELIX  
LOCATION: 333..353  
OTHER INFORMATION: Eisenberg, Schwarz, Komarony, Wall identification method,  
OTHER INFORMATION: helix  
NAME/KEY: MYRISTATE  
LOCATION: 341..345  
OTHER INFORMATION: Prosite match  
NAME/KEY: PHOSPHORYLATION  
LOCATION: 350  
OTHER INFORMATION: potential protein kinase C, Prosite match

US-09-853-526-4

## Query Match

97.0%; Score 353; DB 22; Length 353;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 353; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 12 MYLLPSVVLGTAPTYVLAAGVWRLLSAFLPARFYQALDDRLCYVQSMVLEFFENYTG 71  
DB 1 MYLLPSVVLGTAPTYVLAAGVWRLLSAFLPARFYQALDDRLCYVQSMVLEFFENYTG 60  
QY 72 VOILLYGDLPKNNENIYLANHSTVDWIVADILAIROMALGHVRYVLEKGLKMLPLYGC 131  
|||||

```

Db 61 VOILLVGLDLPKKNENITILANHOSTVDWIVADILAIROMALGHVRYVLEKGLKWLPLYGC 120
; LOCATION: (78)..(78)
; OTHER INFORMATION: PHOSPHORYLATION, potential Tyrosine kinase site, Prosite match
Oy 132 YFAOHGIVYKRSAPFNEKEMRNKLOSVDAGTPMYLVTFPEGTRVNPOTKVLASAOAF 191
; NAME/KEY: MOD_RES
; LOCATION: (84)..(84)
Db 121 YFAOHGIVYKRSAPFNEKEMRNKLOSVDAGTPMYLVTFPEGTRVNPOTKVLASAOAF 180
; OTHER INFORMATION: PHOSPHORYLATION, potential caseline kinase II site, Prosite mat
; NAME/KEY: SITE
; LOCATION: (94)..(115)
Oy 192 AAORGLAVLKHVLTPIKATKHAFCDMKNYLDAYDVTVVYEGKDDGGRRESPTMTEFL 251
; OTHER INFORMATION: Potential leucine zipper site, Prosite match
; NAME/KEY: LIPID
Db 181 AAORGLAVLKHVLTPIKATKHAFCDMKNYLDAYDVTVVYEGKDDGGRRESPTMTEFL 240
; LOCATION: (119)..(123)
; OTHER INFORMATION: MYRISTATE, Prosite match
; NAME/KEY: SITE
Oy 252 CCECPRIHIIHIDRIKDDVPEOEHRRLVHERFEIKDKMLIEFYSPDPERRKRRPGKS 311
; NAME/KEY: MOD_RES
; LOCATION: (133)..(133)
Db 241 CCECPRIHIIHIDRIKDDVPEOEHRRLVHERFEIKDKMLIEFYSPDPERRKRRPGKS 300
; OTHER INFORMATION: PHOSPHORYLATION, potential protein kinase C, Prosite match
; NAME/KEY: MOD_RES
Oy 312 VNSKISIKTLPMSLTLGTLGMLTMDAGRKLYVTWTYGTLLGLWTIKR 364
; LOCATION: (147)..(147)
; OTHER INFORMATION: PHOSPHORYLATION, potential caseline kinase II, Prosite match
; NAME/KEY: MOD_RES
Db 301 VNSKISIKTLPMSLTLGTLGMLTMDAGRKLYVTWTYGTLLGLWTIKR 353
; LOCATION: (194)..(194)
; OTHER INFORMATION: PHOSPHORYLATION, potential protein kinase C, Prosite match
; NAME/KEY: MOD_RES
; LOCATION: (215)..(215)
; OTHER INFORMATION: PHOSPHORYLATION, potential Tyrosine kinase site, Prosite match
; NAME/KEY: MOD_RES
; LOCATION: (221)..(221)
; OTHER INFORMATION: SULFATATION, Prosite match
; NAME/KEY: MOD_RES
; LOCATION: (233)..(233)
; OTHER INFORMATION: PHOSPHORYLATION, potential cAMP and cGMP dependant protein kin
; OTHER INFORMATION: site, Prosite match
; NAME/KEY: MOD_RES
; LOCATION: (235)..(235)
; OTHER INFORMATION: PHOSPHORYLATION, potential caseline kinase II site, Prosite mat
; NAME/KEY: MOD_RES
; LOCATION: (306)..(306)
; OTHER INFORMATION: PHOSPHORYLATION, potential protein kinase C, Prosite match
; NAME/KEY: LIPID
; LOCATION: (310)..(330)
; OTHER INFORMATION: Eisenberg, Schwarz, Komarony, Wall identification method,
; OTHER INFORMATION: potential helix
; NAME/KEY: LIPID
; LOCATION: (319)..(323)
; OTHER INFORMATION: MYRISTATE, Prosite match
; NAME/KEY: LIPID
; LOCATION: (323)..(327)
; OTHER INFORMATION: MYRISTATE, Prosite match
; NAME/KEY: MOD_RES
; LOCATION: (329)..(329)
; OTHER INFORMATION: AMIDATION, Prosite match
; NAME/KEY: HELIX
; LOCATION: (333)..(353)
; OTHER INFORMATION: Eisenberg, Schwarz, Komarony, Wall identification method,
; OTHER INFORMATION: potential helix
; NAME/KEY: LIPID
; LOCATION: (341)..(345)
; OTHER INFORMATION: MYRISTATE, Prosite match
; NAME/KEY: MOD_RES
; LOCATION: (350)..(350)
; OTHER INFORMATION: PHOSPHORYLATION, potential protein kinase C, Prosite match
; US-09-901-484A-4

RESULT 13
US-09-901-484A-4
; Sequence 4, Application US/09901484A
; GENERAL INFORMATION:
; APPLICANT: Cohen, Daniel
; APPLICANT: Blumenfeld, Marta
; APPLICANT: Chumakov, Ilya
; APPLICANT: Bouguetel, Lydie
; TITLE OF INVENTION: Prostate Cancer Gene
; FILE REFERENCE: GEN-T11XC3D2
; CURRENT APPLICATION NUMBER: US/09/901,484A
; PRIOR FILING DATE: 2001-07-09
; PRIOR APPLICATION NUMBER: US 08/996,306
; PRIOR FILING DATE: 1997-12-22
; PRIOR APPLICATION NUMBER: US 60/099,658
; PRIOR FILING DATE: 1998-09-09
; PRIOR APPLICATION NUMBER: US 09/218,207
; PRIOR FILING DATE: 1998-12-22
; PRIOR APPLICATION NUMBER: US 09/338,907
; PRIOR FILING DATE: 1999-06-23
; PRIOR APPLICATION NUMBER: US 09/853,526
; PRIOR FILING DATE: 2001-05-11
; NUMBER OF SEQ ID NOS: 578
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 4
; LENGTH: 353
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: HELIX
; LOCATION: (1)..(33)
; OTHER INFORMATION: Rao and Argos identification method, potential helix
; NAME/KEY: HELIX
; LOCATION: (4)..(20)
; OTHER INFORMATION: Klein, Kanehisa and Delisi identification method, potential helix
; NAME/KEY: HELIX
; LOCATION: (4)..(24)
; OTHER INFORMATION: Eisenberg, Schwarz, Komarony, Wall identification method,
; NAME/KEY: LIPID
; LOCATION: (12)..(16)
; OTHER INFORMATION: MYRISTATE, Prosite match
; NAME/KEY: HELIX
; LOCATION: (50)..(70)
; OTHER INFORMATION: Eisenberg, Schwarz, Komarony, Wall identification method,
; NAME/KEY: CARBOHYD
; LOCATION: (57)..(59)
; OTHER INFORMATION: Prosite match
; NAME/KEY: HELIX
; LOCATION: (76)..(96)
; OTHER INFORMATION: Eisenberg, Schwarz, Komarony, Wall identification method,
; OTHER INFORMATION: potential helix
; NAME/KEY: MOD_RES

```

```

Query Match 97.0%; Score 353; DB 23; Length 353;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 353; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Oy 12 MYLLPSVVLGTAPTYVLAAGVWRLLSAFLPARYQALDDRLCYCYQSMVLFPPENYTG 71
; NAME/KEY: MOD_RES
; LOCATION: (133)..(133)
Db 1 MYLLPSVVLGTAPTYVLAAGVWRLLSAFLPARYQALDDRLCYCYQSMVLFPPENYTG 60
; NAME/KEY: MOD_RES
; LOCATION: (133)..(133)
Oy 72 VOILLVGLDLPKKNENITILANHOSTVDWIVADILAIROMALGHVRYVLEKGLKWLPLYGC 131
; NAME/KEY: MOD_RES
; LOCATION: (133)..(133)
Db 61 VOILLVGLDLPKKNENITILANHOSTVDWIVADILAIROMALGHVRYVLEKGLKWLPLYGC 120
; NAME/KEY: MOD_RES

```

```
QY 132 YFAOHGIIYKRSKAFNEKEMRNKLSQSYVDAGTPMYLVIFPESTRYNPEQTKVLSASQAF 191
|
|
|
Db 121 YFAOHGIIYKRSKAFNEKEMRNKLSQSYVDAGTPMYLVIFPESTRYNPEQTKVLSASQAF 180
|
|
|
QY 192 AAORGLAVLKHVLTLPRIKATHVAFDCMKNTLDAIYDVTYVEGKDDGGRRESPTMTTEFL 251
|
|
|
Db 181 AAORGLAVLKHVLTLPRIKATHVAFDCMKNTLDAIYDVTYVEGKDDGGRRESPTMTTEFL 240
|
|
|
QY 252 CKCEPKIHIIHIDRIKDDVPEEOEHMRMLHERFEIKDKMLIEFYESPDEPERKRRPFGKS 311
|
|
|
Db 241 CKCEPKIHIIHIDRIKDDVPEEOEHMRMLHERFEIKDKMLIEFYESPDEPERKRRPFGKS 300
|
|
|
QY 312 VNSKLSIKKTLPMSMLISGLTAGMLMTDAGRKLVTWMTWYIGTLGCLMWYTIKA 364
|
|
|
Db 301 VNSKLSIKKTLPMSMLISGLTAGMLMTDAGRKLVTWMTWYIGTLGCLMWYTIKA 353
```

```
RESULT 14
US-09-915-582-59
; Sequence 59, Application US/09915582
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: 17 Human Secreted Proteins
; FILE REFERENCE: PS723P1
; CURRENT APPLICATION NUMBER: US/09/915,582
; CURRENT FILING DATE: 2001-07-27
; PRIOR APPLICATION NUMBER: PCT/US01/01431
; PRIOR FILING DATE: 2001-01-17
; PRIOR APPLICATION NUMBER: 60/179,065
; PRIOR FILING DATE: 2000-01-31
; PRIOR APPLICATION NUMBER: 60/180,628
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: 60/231,968
; PRIOR FILING DATE: 2000-09-12
; NUMBER OF SEQ ID NOS: 97
; SOFTWARE: Patentln Ver. 2.0
; SEQ ID NO 59
; LENGTH: 353
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-915-582-59
```

```
Query Match 97.0%; Score 353; DB 23; Length 353;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 353; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 12 MRYLLPSVVLGTAPTYVLAMGWRLLSAFLPARFYQALDDRLCYQOSWVLEFFENYTG 71
|
|
|
Db 1 MRYLLPSVVLGTAPTYVLAMGWRLLSAFLPARFYQALDDRLCYQOSWVLEFFENYTG 60
|
|
|
QY 72 VOILLGDLPRKKNENIITYLANHOSTVDWIVADILAIROMALGHVRVLEKGLMPLPYGC 131
|
|
|
Db 61 VOILLGDLPRKKNENIITYLANHOSTVDWIVADILAIROMALGHVRVLEKGLMPLPYGC 120
|
|
|
QY 132 YFAOHGIIYKRSKAFNEKEMRNKLSQSYVDAGTPMYLVIFPESTRYNPEQTKVLSASQAF 191
|
|
|
Db 121 YFAOHGIIYKRSKAFNEKEMRNKLSQSYVDAGTPMYLVIFPESTRYNPEQTKVLSASQAF 180
|
|
|
QY 192 AAORGLAVLKHVLTLPRIKATHVAFDCMKNTLDAIYDVTYVEGKDDGGRRESPTMTTEFL 251
|
|
|
Db 181 AAORGLAVLKHVLTLPRIKATHVAFDCMKNTLDAIYDVTYVEGKDDGGRRESPTMTTEFL 240
|
|
|
QY 252 CKCEPKIHIIHIDRIKDDVPEEOEHMRMLHERFEIKDKMLIEFYESPDEPERKRRPFGKS 311
|
|
|
Db 241 CKCEPKIHIIHIDRIKDDVPEEOEHMRMLHERFEIKDKMLIEFYESPDEPERKRRPFGKS 300
|
|
|
QY 312 VNSKLSIKKTLPMSMLISGLTAGMLMTDAGRKLVTWMTWYIGTLGCLMWYTIKA 364
|
|
|
Db 301 VNSKLSIKKTLPMSMLISGLTAGMLMTDAGRKLVTWMTWYIGTLGCLMWYTIKA 353
```

```
RESULT 15
US-09-915-582-79
; Sequence 79, Application US/09915582
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: 17 Human Secreted Proteins
; FILE REFERENCE: PS723P1
; CURRENT APPLICATION NUMBER: US/09/915,582
; CURRENT FILING DATE: 2001-07-27
; PRIOR APPLICATION NUMBER: PCT/US01/01431
; PRIOR FILING DATE: 2001-01-17
; PRIOR APPLICATION NUMBER: 60/179,065
; PRIOR FILING DATE: 2000-01-31
; PRIOR APPLICATION NUMBER: 60/180,628
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: 60/231,968
; PRIOR FILING DATE: 2000-09-12
; NUMBER OF SEQ ID NOS: 97
; SOFTWARE: Patentln Ver. 2.0
; SEQ ID NO 79
; LENGTH: 353
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-915-582-79
```

```
Query Match 97.0%; Score 353; DB 23; Length 353;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 353; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 12 MRYLLPSVVLGTAPTYVLAMGWRLLSAFLPARFYQALDDRLCYQOSWVLEFFENYTG 71
|
|
|
Db 1 MRYLLPSVVLGTAPTYVLAMGWRLLSAFLPARFYQALDDRLCYQOSWVLEFFENYTG 60
|
|
|
QY 72 VOILLGDLPRKKNENIITYLANHOSTVDWIVADILAIROMALGHVRVLEKGLMPLPYGC 131
|
|
|
Db 61 VOILLGDLPRKKNENIITYLANHOSTVDWIVADILAIROMALGHVRVLEKGLMPLPYGC 120
|
|
|
QY 132 YFAOHGIIYKRSKAFNEKEMRNKLSQSYVDAGTPMYLVIFPESTRYNPEQTKVLSASQAF 191
|
|
|
Db 121 YFAOHGIIYKRSKAFNEKEMRNKLSQSYVDAGTPMYLVIFPESTRYNPEQTKVLSASQAF 180
|
|
|
QY 192 AAORGLAVLKHVLTLPRIKATHVAFDCMKNTLDAIYDVTYVEGKDDGGRRESPTMTTEFL 251
|
|
|
Db 181 AAORGLAVLKHVLTLPRIKATHVAFDCMKNTLDAIYDVTYVEGKDDGGRRESPTMTTEFL 240
|
|
|
QY 252 CKCEPKIHIIHIDRIKDDVPEEOEHMRMLHERFEIKDKMLIEFYESPDEPERKRRPFGKS 311
|
|
|
Db 241 CKCEPKIHIIHIDRIKDDVPEEOEHMRMLHERFEIKDKMLIEFYESPDEPERKRRPFGKS 300
|
|
|
QY 312 VNSKLSIKKTLPMSMLISGLTAGMLMTDAGRKLVTWMTWYIGTLGCLMWYTIKA 364
|
|
|
Db 301 VNSKLSIKKTLPMSMLISGLTAGMLMTDAGRKLVTWMTWYIGTLGCLMWYTIKA 353
```

Search completed: August 28, 2002, 11:25:19  
Job time: 497 sec